ENGINEERING OF AN EFFICIENT AND ENANTIOSELECTIVE BIOCATALYST FOR THE PREPARATION OF CHIRAL PHARMACEUTICAL INTERMEDIATES

BY

WENG LIN TANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign,
in the program jointly administered with the
National University of Singapore, 2011

Doctoral Committee:

Professor Huimin Zhao, Chair
Associate Professor Zhi Li, NUS, Co-Director of Research
Associate Professor Shing Bor Chen, NUS
Associate Professor Christopher V. Rao
Assistant Professor Rudiyanto Gunawan, ETH Zürich
Professor Daniel W. Pack

Abstract

This Ph.D. thesis focuses on the engineering of an efficient and enantioselective biocatalyst via direct evolution and genetic engineering for the enantioselective hydroxylation of non-activated carbon atom, a useful but challenging reaction for the synthesis of chiral pharmaceutical intermediates. Our target enzyme is the novel P450pyr enzyme from *Sphingomonas* sp. HXN-200 that was found to catalyze the regio- and stereoselective hydroxylation of non-activated carbon atom with broad substrate range, high activity, excellent regioselectivity, and good to excellent enantioselectivity. Our target reaction is the enzymatic hydroxylation of *N*-benzyl pyrrolidine to its corresponding (*R*)- and (*S*)-*N*-benzyl-3-hydroxypyrrolidines which are important pharmaceutical intermediates.

In this thesis, a two-enzyme-based colorimetric high-throughput *ee* screening assay and a mass spectrometry-based high-throughput *ee* screening assay were developed. The P450pyr monooxygenase was engineered by directed evolution for the enantioselective hydroxylation of *N*-benzyl pyrrolidine. Several mutants exhibiting increased and/or inverted enantioselectivity were identified, with product *ee* of 83% (*R*) and 65% (*S*) for mutants 1AF4A and 11BB12, respectively. The wild type P450pyr and its mutants were also purified and reconstituted with their auxiliary electron transport proteins, ferredoxin and ferredoxin reductase *in vitro*. The mutants were then used to catalyze the hydroxylations of a range of different substrates using whole-cell assays to investigate the changes in product *ee*. In addition, an efficient biocatalytic system with cofactor recycling was developed by co-expressing a glucose dehydrogenase from *Bacillus substilis* or a phosphite

dehydrogenase from *Pseudomonas stutzeri* together with the P450pyr system in a recombinant *Escherichia coli*.

To Papa, Mama, Jun Jun

and Pippo

Acknowledgements

This Ph.D. thesis would not have been possible without my advisors, Professor Huimin Zhao and Associate Professor Zhi Li, whose constant guidance, great patience and understanding have led me through to the completion of my graduate career. Many thanks to my thesis examiners, Associate Professor Shing Bor Chen (Chair), Professor Wilfred Chen, Associate Professor Christopher V. Rao and Assistant Professor Rudiyanto Gunawan, for their valued support and suggestions.

In particular, I am indebted to Dr. Sheryl Rubin Pitel, my mentor, who taught me the basics of molecular biology and how to conduct high quality research. I would also like to thank Dr. Ryan Sullivan, Dr. Nikhil Nair, Dr. Yoo-Seong Choi, Dr. Zengyi Shao, Dr. Michael McLachlan and Dr. Zunsheng Wang for their helpful discussions and extremely useful suggestions. A big thank you to Carl Denard, Luigi Chanco, Ryan Cobb, Ning Sun, Dr. Byoungjin Kim, Liang Xue, Wei Zhang, Wen Wang, Quang Son Pham and all the current and former members of Prof. Huimin Zhao's and Prof. Zhi Li's laboratory for their wonderful friendship and for making my Ph.D. life interesting and wonderful. Special thanks to Dr. Yongzheng Chen who worked with me on the high-throughput mass spectrometry-based assay and for helping me to synthesize various chemical compounds for my biocatalysis work.

Lastly but most importantly, I would like to thank my family for their love, support and encouragement. Everything that I have achieved today would not have been possible without them.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

Amp	ampicillin
BSA	bovine serum albumin
cdw	cell dry weight
Cm	chloramphenicol
epPCR	error-prone polymerase chain reaction
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high-performance liquid chromatography
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl-β-D-thiogalactospyranoside
Kan	kanamycin
LB	Luria-Bertani
MS	mass spectrometry
LC-MS	liquid chromatography-mass spectrometry
NAD^+	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
$NADP^+$	nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate
NBT	nitroblue tetrazolium
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMS	phenazine methosulfate
PTDH	phosphite dehydrogenase
Sm	streptomycin/spectinomycin

SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TB	Terrific broth
WT	wild type

Amino Acid Abbreviations

A	alanine
C	cysteine
D	aspartate
E	glutamate
F	phenylalanine
G	glycine
H	histidine
I	isoleucine
K	lysine
L	leucine
M	methionine
N	asparagine
P	proline
Q	glutamine
R	arginine
S	serine
T	threonine
V	valine
W	tryptophan
Y	tyrosine

Chapter 1 : Introduction

1.1 Industrial Biotechnology

Industrial biotechnology, also known as white biotechnology, is the application of modern biotechnology to the sustainable production of chemicals (which includes fine chemicals and pharmaceuticals), materials, and fuels from renewable sources, using living cells and/or their enzymes. This field is widely regarded as the third wave of biotechnology, distinct from the first two waves (medical or red biotechnology and agricultural or green biotechnology). Much interest has been generated in this field mainly because industrial biotechnology is often associated with reduced energy consumption, greenhouse gas emissions, and waste generation, and also may enable the paradigm shift from fossil fuel-based to bio-based production of value-added chemicals.

The fundamental force that drives the development and implementation of industrial biotechnology is the market economy, as biotechnology promises highly efficient processes at lower operating and capital expenditures. In addition, political and societal demands for sustainability and environment-friendly industrial production systems, coupled with the depletion of crude oil reserves and a growing world demand for raw materials and energy, will continue to drive this trend forward.[1] McKinsey & Company predicted that in 2010 industrial biotechnology will account for 10 percent of sales within the chemical industry, amounting to US$125 billion in value (http://www.chemie.de/news/e/pdf/news_chemie.de_56388.pdf). In the US, bio-based pharmaceuticals account for the largest share of the biotechnology market followed by bio-ethanol, other bio-based chemicals, and bio-diesel.[2] Other major

players in industrial biotechnology include the European Union,[3,4] China, India, and Brazil. In China alone, the value of bio-based chemical products exceeded US$60.5 billion in 2007.[5]

Government policies including tax incentives, mandatory-use regulations, research and development, commercialization support, loan guarantees, and agricultural feedstock support programs have helped fuel the adoption of industrial biotechnology. Moreover, breakthroughs in enzyme engineering, metabolic engineering, synthetic biology, and the expanding 'omics' toolbox coupled with computational systems biology are expected to speed up industrial application of biotechnology. These advances have provided scientists with toolsets to engineer enzymes and whole-cells, by expanding the means to identify, understand and make perturbations to the complex machinery within the microorganisms.

1.2 Chemo-, Regio- and Enantioselective Biocatalysis

Biocatalysis is one of the oldest chemical transformations known to humans, with the oldest records of brewing dating back to about 6000 years ago. The employment of enzymes and whole-cells in chemical, pharmaceutical, as well as food, textile and paper industries is increasing rapidly. Whole-cells are often used in reactions which require the regeneration of expensive cofactors, whereas isolated enzymes are typically used for hydrolytic or isomerization reactions. Nevertheless, in industrial settings, whole-cells are usually preferred over purified enzymes, even for single step transformations, to avoid enzyme purification costs.

Biocatalysis offers several significant advantages over chemocatalysts. Firstly, since enzymes function at moderate temperatures and pressures, they require less

energy input. Secondly, enzymes can be extremely fast, increasing reaction rates by 4 to 12 orders of magnitude. In some cases, enzymes can be so fast that their turnover rates are limited only by diffusion of substrate(s) and product(s) to and from the active site. Thirdly, since biocatalysts are biodegradable and rarely contain heavy metals, waste streams are much more manageable, requiring minimal treatment. Fourthly, since most biocatalysts function in aqueous media, effluent volatile organic compound (VOC) emissions from production plants can be drastically reduced. Lastly but most importantly, enzymes can be highly selective (regio-, diastereo-, enantio-, or chemo-), and thus are able to replace multistep reactions or difficult purification schemes. This can circumvent the need for many blocking and deblocking steps that are required for stereo- or regio-selective reactions.

The high selectivities of enzyme make them attractive for industrial applications due to its increased product concentrations and productivities besides having no undesirable by-products. Despite the considerable progress in chiral synthesis, organic chemists are still struggling with the complexity of using traditional organic chemistry to synthesize chiral compounds. Thus, the focus is now switched to biocatalysis - a simple yet efficient method to synthesize enantiopure compounds as chiral building blocks for drugs and agrochemicals. The use of biocatalysis to produce enantiopure compounds can be divided into two approaches: (1) kinetic resolution of a racemic mixture where the enzyme is used to convert one of the enantiomers at a higher reaction rate than the other enantiomer; and (2) asymmetric biocatalysis starting with a prochiral substrate to produce different enantiomers in different quantities.

Chirality is a very important factor in the efficacy of many drug products. A survey of major pharmaceutical companies such as GlaxoSmithKline (UK), AstraZeneca (UK) and Pfizer (US) showed that more than half of the drug compounds examined contained one or more chiral centers.[6] Furthermore, an enantiomeric purity of at least 99.5% is often necessary to meet regulatory requirements. Selectivities of >95% are difficult to achieve by chemocatalysis and, if essential to synthesis, would require the use of biocatalysis. At present, 22 of 38 large-scale asymmetric syntheses already incorporated biocatalysis.[7] Recognizing its importance, many fine chemical and pharmaceutical companies have started to focus on acquiring biocatalysis expertise (Table 1.1), and those that have already done so are trying to maintain their positions as technological leaders.

Table 1.1. Biotransformations developed by the pharmaceutical industry.[8]

Target compound	**Company**	**Reaction**	**Biocatalyst**
SCH56592	Schering Plough	Acylation	CALB
β-Lactams	Glaxo Smith Kline	Acylation	CALB
Lotrafiban	Glaxo Smith Kline	Hydrolysis	CALB
Paclitaxel	Bristol-Myers Squibb	Hydrolysis	*Pseudomonas lipase* AK *P. cepacia* lipase PS-30
HMG-CoA reductase inhibitor	Bristol-Myers Squibb	Acylation	*P. cepacia* lipase PS-30
SCH66336	Schering Plough	Acylation	*P. aeruginosa* lipase
Xemilofiban	Monsanto	Hydrolysis	*E. coli* penicillin acylase
Renin inhibitor	Hoffmann La Roche	Hydrolysis	B. licheniformis subtilisin
Lamivudine	Glaxo Smith Kline	Hydrolysis	*E. coli* cytidine deaminase
AG7088	Pfizer	Reduction	*L. mesenteroides* D-LDH *C. boidinii* FDH

Table 1.1 (cont.)

ACE inhibitor	Ciba-Geigy	Reduction	*S. epidermis* D-LDH
LY300164	Eli Lilly	Reduction	*Z. rouxii* dehydrogenase
Omapatrilat	Bristol-Myers Squibb	Reductive amination	*T. intermedius* PDH
N-Butyl DNJ	Pharmacia	Oxidation	*G. oxydans* SDH
2-Quinoxaline-carboxylic acid	Pfizer	Oxidation	*A. repens* MO
HMG-CoA reductase inhibitor	Merck Sharp and Dohme	Oxidation	*N. autotropica* MO
Lobucavir prodrug	Bristol-Myers Squibb	Regioselective acylation	*B. licheniformis* subtilisin

CALB: *Candida antartica* lipase, FDH: formate dehydorgenase, D-LDH: D-lactate dehydrogenase, MO: monooxygenase, PDH: phenylalanine dehydrogenase, SDH: sorbitol dehydrogenase.

1.3 Enzymatic Hydroxylation of Non-Activated Hydrocarbons

Regio- and stereoselective hydroxylations of non-activated carbon atoms represent a significant challenge in classical organic chemistry.[9] However, this reaction can be carried out via biocatalysis with monooxygenases using molecular oxygen as oxidant. Several monooxygenases such as cytochrome P450cam, P450BM-3, methane monooxygenases (MMO) and membrane-bound alkane hydroxylase (alkB) have been extensively investigated. These biohydroxylations are very useful for the preparation of enantiopure compounds that are useful pharmaceutical intermediates.

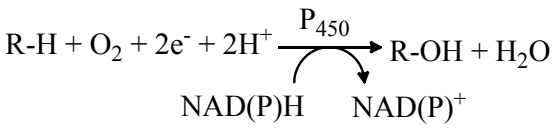
1.3.1 Cytochrome P450 Monooxygenase

First discovered about 50 years ago, cytochrome P450 enzymes constitute the largest superfamily of heme-containing monooxygenases (E.C.1.14.-.-) that have the ability to oxidize a broad range of substrates, often at non-reactive carbon centers.[10] The name cytochrome P450 is derived from the following: cytochrome stands for a hemoprotein, P for pigment, and 450 reflects the absorption peak of the reduced CO-

bound complex at 450 nm. Cytochrome P450s are widely distributed in prokaryotes and eukaryotes, with over 11,000 P450s identified to date (http://drnelson.uthsc.edu/CytochromeP450.htmL). The sequence of the P450 polypeptide chain can be quite diverse with P450s from the same family differing by up to 60% in their amino acid sequence.

P450 enzymes can be divided into four classes, depending on their redox partners. Class I P450s include most bacterial monooxygenases that require a FAD-containing reductase and an iron sulfur redoxin (P450cam from *Pseudomonas putida*). Class II enzymes require only a FAD/FMN-containing P450 reductase for the transfer of electrons and are found mostly attached to the endoplasmatic reticulum. P450BM-3 from *Bacillus megaterium*, CYP102A2 from *Bacillus subtilis* and CYP505from *Fusarium oxysporum* belong to this class of enzymes. They are self-sufficient as they contain the P450 monooxygenase and the reductase domain on a single peptide chain. Class III enzymes are self-sufficient and require no electron donor as they convert peroxygenated substrates that already contain oxygen. Meanwhile, P450s from class IV receive electrons directly from NAD(P)H (P450nor from *F. oxysporum*).

The most common hydroxylation reaction catalyzed by cytochrome P450 is the insertion of one atom of oxygen into a nonactivated carbon atom of an organic substrate (RH), which is not possible by standard chemical methods, while the other oxygen atom is reduced to water.



Scheme 1.1. General hydroxylation reaction catalyzed by P450 monooxygenases.

The catalytic turnover of cytochrome P450 begins with the binding of substrate followed by the introduction of the first electron from NADPH via an electron transfer chain.[11] Next, the oxygen binds and accepts a second electron to produce a ferric peroxy anion. This anion is protonated to form the ferric hydroperoxy complex, which is subjected to a heterolytic cleavage with the formation of a putative ferryl species, equivalent to a Fe(V)=O species. This reactive electrophilic iron-oxo intermediate then attacks the substrate to yield the hydroxylated product, which dissociates to allow the cycle to begin again.

Cytochrome P450s are involved in the biotransformation of drugs, the bioconversion of xenobiotics, the metabolism of chemical carcinogens, the biosynthesis of physiologically important compounds such as steroids, fatty acids, eicosanoids, fat-soluble vitamins, bile acids, the conversion of alkanes, terpenes and aromatic compounds as well as the degradation of herbicides and insecticides.[11] P450s are involved in diverse reactions including hydroxylation, *N*-, *O*-, *S*-dealkylation, sulphoxidation, epoxidation, deamination, desulphuration, dehalogenation, peroxidation and *N*-oxide reduction.[12] P450s can also catalyze multi-step reactions. One example is the transformation of artemisinic acid, a precursor of the antimalarial drug artemisin, by CYP71AV1.[13] In this example, amorpha-4,11-diene, a natural product of *E. coli* or *Saccharomyces cerevisiae*, was initially hydroxylated twice by CYP71AV1. The hydroxylation product was spontaneously dehydrated to form an aldehyde which was subsequently converted by CYP71AV1 to form artemisinic acid. In another example, $P450_{scc}$ was used to convert cholesterol by two successive hydroxylation reactions into (20*R*, 22*R*),20,22-dihydroxycholesterol. The $P450_{scc}$ was then used to catalyze the C-C bond cleavage to yield prenenolone and isocaproic

acid.[14] Despite the large application scope of P450 enzymes as biocatalysts, they are rarely implemented in industrial processes mainly due to their instability, complexity and often low catalytic turnover.[15] Nevertheless, with the increasing advancement in the powerful methodologies of genetic engineering, P450 enzymes can now be engineered with tailor-made properties, such as the engineering of recombinant expression, physical properties, catalytic activity and interactions with redox partners.[16]

1.3.2 Methane Monooxygenases

Methane monooxygenase (MMO) belongs to the class of oxidoreductase enzymes (EC 1.14.13.25) and is able to oxidize the C-H bond in methane as well as other alkanes. There are two well-investigated forms of MMO: the soluble form (sMMO) and the particulate form (pMMO). The active site in sMMO contains a di-iron center bridged by an oxygen atom (Fe-O-Fe), whereas the active site in pMMO utilizes copper. Over the decades, intensive investigation of these hydroxylation reactions have been carried out either with the methanogenic microorganisms or purified enzymes.[17,18] Methane monooxygenases (MMOs) isolated from *Methylococcus capsulatus* Bath and *Methylosinus trichosporium* OB3b are found to be able to hydroxylate simple alkanes with NADH as cofactor.[19] The MMO from Methylococcus capsulatus is able to hydroxylate a series of simple alkanes to alcohols, such as *n*-pentane to (*R*)-2-pentanol, although the yields are significantly reduced with increasing chain lengths in the substrates.[17]

1.3.3 Membrane-bound Alkane Hydroxylase (AlkB)

AlkB is an integral membrane non-heme iron protein that was first discovered in a hexane-degrading pseudomonad known as *Pseudomonas putida* GPo1.[20] The AlkB system consists of three components: the AlkB monooxygenase, rubredoxin and the electron-providing NADH-dependent flavoprotein rubredoxin reductase. The AlkB enzymes are very diverse and have different substrate ranges with the capability to oxidize C_5 to C_{16} *n*-alkanes, as well as branched, cyclic aliphatic and aromatic compounds.[21] Extensive work has been carried out to understand the structure-function relationship of this class of enzyme. One interesting example is a protein engineering study where a tryptophan residue located in the middle of one of six transmembrane helices was found to limit the length of the alkane substrates.[22]

1.4 Protein Engineering

One of the most important tools for industrial biotechnology is protein engineering. More often than not, a wild-type enzyme discovered in nature is not suitable for an industrial process. There is a need to engineer and optimize enzyme performance in terms of activity, selectivity on non-natural substrates, thermostability, tolerance towards organic solvents, enantioselectivity, and substrate/product inhibition among others in order for the enzymatic process to be commercially viable[23] (Figure 1.1). There are two general approaches for protein engineering: rational design and directed evolution.

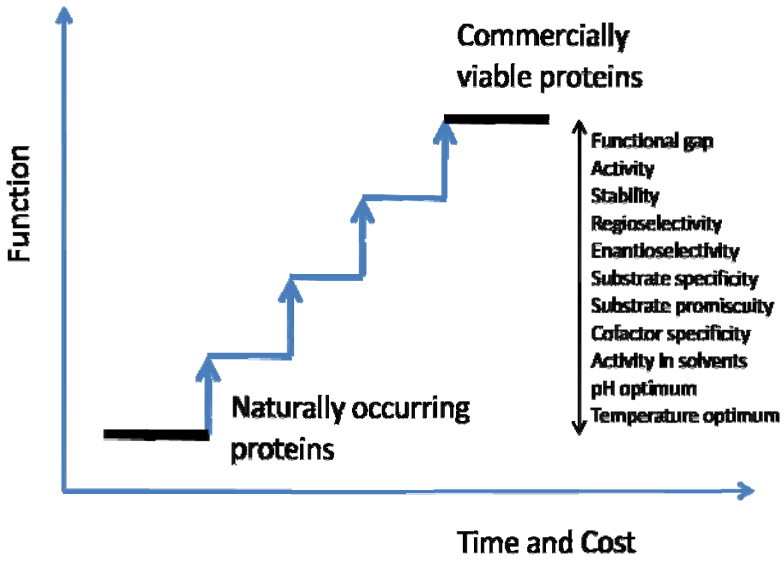


Figure 1.1 A functional gap that exists between the naturally occurring enzymes and the commercially viable enzymes needs to be bridged.

1.4.1 Rational Design

In rational design, the structure, function and catalytic mechanism of the protein must be well understood in order to make desired changes via site-directed mutagenesis. However, such understanding is lacking for most proteins of interest. In addition, although computational protein design algorithms were developed to predict optimal mutations at specific residue positions in the protein, only limited success has been demonstrated.[24-26] One interesting example is in the site-directed mutagenesis of a specific phenylalanine residue (F87) within the hydrophobic channel connecting the surface and the heme-bound active site of CYP102 to enhance its substrate specificity and regioselectivity in fatty acid hydroxylation.[27] By replacing the F87 with different residues, the activity for a variety of aromatic and phenolic compounds was also increased.[28] Based on the crystal structure of cytochrome P450

2C5, Kumar and colleagues have also rationally engineered the P450 2B1 to alter the regioselectivity of progesterone metabolism from 16α- to 21-hydroxylation.[29]

1.4.2 Directed Evolution

Directed evolution is a powerful tool for tailoring enzyme functions in order to close the functional gap between the naturally occurring protein and one that is commercially viable.[30-33] Inspired by natural Darwinian evolution, directed evolution involves creation of genetic diversity via methods such as error-prone polymerase chain reaction (epPCR), saturation mutagenesis, DNA shuffling[34] and Staggered Extension Process (StEP)[35], followed by screening and/or selection for mutants with desired features. Compared to the natural evolution of species which may take millions of years, directed evolution can be carried out within months and with a greater number of parents. Hence, directed evolution is a fast way to develop biocatalysts which have desired characteristics. The advantage of directed evolution over rational design is that it does not need any structural or mechanistic information of the protein of interest and can be carried out with just the knowledge of the gene sequence.

For example, error-prone PCR and site saturation mutagenesis have been used to engineer the activity and regioselectivity of the cytochrome P450 BM-3[36]. Iterative site-specific saturation mutagenesis has also been used to alter the ligand-binding specificity of the human estrogen receptor α (hERα) to recognize non-steroidal synthetic compounds[37-39] and xylose-specific xylose reductase for xylitol synthesis.[40] In addition, a family shuffling approach was used to increase the catalytic activity and thermostability of a type III polyketide synthase, PhlD from the soil bacterium

Pseudomonas fluorescens Pf-5.[41] A summary of directed evolution techniques is shown in Table 1.2.

Table 1.2. Summary of the advantages and disadvantages of selected directed evolution methods (Adapted from Rubin-Pitel and coworkers[42]).

Technique	Advantages	Disadvantages
epPCR	Simplicity Tunable mutation rate	Biased mutagenesis
SeSaM	Unbiased mutagenesis Codon randomization possible	2-3 days to perform Several steps, reagents & enzymes required Special primers required Several purification steps involved
RID	Random insertions and deletion Large diversity possible Codon randomization possible	Several steps, reagents & enzymes required Frameshift mutations possible
RAISE	Random insertions and deletion Codon randomization possible	Frameshift mutations possible DNaseI digestion bias
DNA Shuffling	Robust, flexible Back-crossing to parent removes non-essential mutations Synergistic/additive mutations can be found	DNaseI digestion bias Biased to crossovers in high homology regions Low crossover rate High percentage of parent
Family Shuffling	Exploits natural diversity Accelerated phenotype improvement	DNaseI digestion bias Biased to crossover in high homology regions Need high sequence homology in family Low crossover rate High percentage of parent
RACHITT	No parent genes in shuffled library Higher rate of recombination Recombine genes of low sequence homology	Several steps, reagents & enzymes required Requires synthesis and fragmentation of single-stranded complement DNA
NExT DNA Shuffling	Predictable fragmentation pattern	Non-random fragmentation Several steps, reagents & enzymes required Toxic piperidine used
StEP	Simplicity	Need high homology Low crossover rate Need tight control of PCR
CLERY	Not limited by ligation efficiency of gene into vector	Transformants contain more than one mutant, so rescue and retransformation required

Table 1.2 (cont.)

		Long PCR program for reassembly DNaseI digestion bias Background mutation in plasmid possible Limited diversity
ITCHY	Eliminates recombination bias Structural knowledge not needed Completely homology-independent	Limited to two parents One crossover per iteration Significant fraction of progeny out-of-frame Complex, labor-intensive Single crossovers
SCRATCHY	Eliminates recombination bias Structural knowledge not needed Multiple crossovers possible	Limited to two parents Significant fraction of progeny out-of-frame Complex, labor-intensive DNaseI digestion bias

1.4.2.1 Directed Evolution of P450 Monooxygenases

By combining error-prone PCR (epPCR) and *in vitro* recombination, Glieder et al. have evolved a converted the P450 BM-3 from a fatty acid monooxygenase into one that can hydroxylate hexane and other alkanes with high activity.[43] In fact, the hydroxylation turnover rates of all the liquid alkanes exceed those of the wild type P450 BM-3. The improved mutant enzyme contains 11 amino acid substitutions with only one mutation that is in direct contact with the substrate. The work did not stop there as the P450 BM-3 was further evolved by many rounds of DNA shuffling and recombination, site saturation mutagenesis, site directed mutagenesis, and random mutagenesis in order to finally obtain a P450 propane monooxygenase (P450 PMO).[44] This newly evolved enzyme was shown to have a ~9000-fold increase in k_{cat}/K_M. This demonstrates that directed evolution can be used to completely respecialize the cytochrome P450 for function on a nonnative substrate. In another directed evolution experiment, the regioselectivity of CYP102A3 from *Bacillus substilis* was evolved to

hydroxylate substrates at different subterminals, and especially at terminal carbon chain positions.[45]

One main problem in utilizing enzymes at industrial level is its poor tolerance towards organic cosolvents, which is often necessary especially when using high substrate concentrations. To address this issue, Wong et al. used epPCR and saturation mutagenesis to introduce mutations that significantly improve the activity of P450 BM03 and its F87A mutant in the presence of acetone, acetonitrile, dimethylformamide (DMF), dimethylsulfoxide (DMSO), ethanol, and tetrahydrofuran (THF).[46] In this work, mutants containing a small alanine residue at position 87 have been found to have lower organic solvent tolerance compared to variants with phenylalanine at the same position, as the bulky phenylalanine lies perpendicular to the catalytic heme center and is likely to restrict the access of organic cosolvent to the heme.

Many industrial applications require enzymes that retain their function at elevated temperatures. Arnold and colleagues has reported the evolution of a thermostable P450 BM-3 peroxygenase 21B3 resulting in a mutant which is more thermostable than the wild type P450 BM-3, but retains most of the peroxygenase activity of 21B3.[47]

1.4.3 Screening and Selection

Often, finding an enzyme with desirable properties in a library of mutants generated by directed evolution is akin to looking for a needle in a haystack. Over the past several years, a multitude of screening and/or selection techniques have been developed to isolate the variants of interest. An example of a selection method was

described by Boersma and coworkers in the directed evolution of *B. subtilis* lipase A variants with inverted and improved enantioselectivity.[48] The method is based on the use of an *Escherichia coli* aspartate auxotroph, the growth of which is dependent upon hydrolysis of an enantiomerically pure aspartate ester by desired lipase variants. A covalently binding phosphonate ester of the opposite enantiomer was used as a suicide inhibitor to inactivate less enantioselective variants.

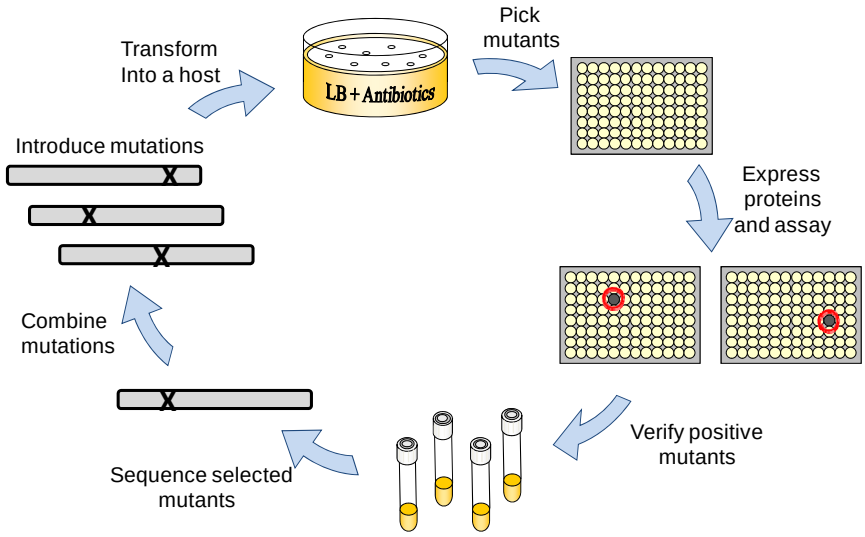


Figure 1.2. A typical screening procedure in a 96-well microtiter plate format

Another commonly used method is microtiter plate-based screening. A typical screening procedure in a 96-well microtiter plate format begins with the generation of a library of mutants which are picked and grown in 96-well plates. The proteins of interest are expressed and are often subjected to a high-throughput assay based on UV-absorption, fluorescence or colorimetric methods. Mutants displaying desired characteristics are then verified and sequenced. The best mutant is then selected as the template for the next round of mutagenesis. The process is repeated in an iterative manner until the goal is achieved or no further improvements are possible (Figure 1.2).

Other screening/selection methods include the agar plate screen, cell-in-droplet screen, cell as microreactor, cell surface display, and *in vitro* compartmentalization, which has been described in earlier reviews.[49,50] Despite the availability of a wide range of screening or selection tools, their applicability is often specific only to a particular substrate/enzyme combination and much effort is still required to customize and optimize a screening/selection method for different directed evolution approaches.

1.4.3.1 High-Throughput Screening Methods for Assaying Enantioselective Enzymes and Biocatalysts

High-throughput screening methods for assaying the enantioselectivity of enzymes can be divided into three categories: (1) assays used for biotransformation from a racemic substrate, (2) assays used for asymmetric catalysis from prochiral substrate, and (3) assays used to quantify the enantiomeric products from an enzymatic reaction.

Assays for enantioselective biotransformation from racemic substrate

Enantioselective enzymatic biotransformation in which an enzyme selectively converts one of the enantiomers of a racemic substrate to product, or also known as kinetic resolution, is the most common basis used in the development of a high-throughput assay. In assays where the ratio of reaction rates of enantiomeric substrates is used to predict the enantiomeric ratio E of a catalyst, the condition is the assay concentration lies below the Michaelis-Menten constant K_M for each enantiomer i.e. very dilute substrate concentration.

One of the earliest high-throughput UV/Vis-based method was an assay developed for the lipase-catalyzed kinetic resolution of chiral *p*-nitrophenol esters, with the goal of evolving highly enantioselective mutants of the lipase from *Pseudomonas aeruginosa*.[51] Lipase-catalyzed hydrolysis generates *p*-nitrophenolate which shows a strong UV/Vis absorption at 405nm. The (*R*)- and the (*S*)-esters were tested separately pairwise on a 96-well microtiter plate using a simple UV/Vis-based plate reader, enabling 48 mutants to be screened in a few minutes. This system has a few drawbacks such as the need for pure enantiomers, requirement of a built-in chromophore (*p*-nitrophenol), whose presence in the substrate may misdirect the results in a directed evolution experiment, and also the need for clear solutions for spectrophotometric measurements. Also, since the (*S*)- and (*R*)-substrates are tested separately pairwise, the competitive binding of enantiomers to the enzyme has been ignored. This may lead to erroneous results in the estimated enantiomeric ratio, E (errors up to ±70%) as the contribution of K_M to the overall selectivity of the enzyme has been ignored.

To minimize this error, the Quick E method made use of a reference compound resorufin tetradecanoate to introduce some competitive binding in the enzymatic reaction.[52] This method allows the (*S*)- and (*R*)-substrates to compete against each other indirectly by competing each enantiomer against a reference compound. The pH indicator assay, on the other hand,[53] utilizes the indicator color change which correlates to the number of proton released during the hydrolysis of esters such that the need for a built-in chromophore is obviated.

Assays based on fluorescence to screen for enantioselectivity is attractive due to its high sensitivity (allows for use of very diluted substrates or small amounts of

catalysts), simplicity and possibility of real-time analysis. This has been demonstrated by Tumambac and Wolf using a C_2-symmetric fluorosensor (1,8-diquinolylnaphthalene *N,N'*-dioxide) employed in real-time enantioselective analysis of the enzymatic kinetic resolution of racemic trans-1,2-diaminocyclohexane.[54] Adding the fluorosensor to (*R,R*)-**3** significantly enhances fluorescence of the sensor, whereas the (*S,S*)-enantiomer has little effect. The fluorescence intensity was not affected by the addition or presence of other analytes such as bisamidoester and monoamidoester, implying that this screening method do not require tedious purification and derivatization steps. In the course of the reaction, as the fluorescence intensity decreased due to the decreasing relative amount of (*R,R*)-**3**, the enantiomeric excess, *ee* of the rac-**3** can be monitored real-time.

Hwang and Kim demonstrated a Cu(II) amine complex formation method to measure the apparent enantioselectivity (E_{app}) of ω-transaminase by measuring reaction rates of pure enantiomers (*R*)- and (*S*)-aromatic amines respectively.[55] The product α-amino acids will form a blue complex with the Cu(II) ion, which is quantifiable using UV/Vis spectrophotometer. This method still requires some fine tuning as the E_{app} value shows large errors when compared to the actual E value by HPLC analysis due to the neglect of the competitive binding of the two enantiomers to the enzyme.

The most direct method of screening would be to link improved enzyme activity to the survival or growth rates of cells which expresses the enzyme. Reetz and Rüggeberg demonstrated a differential growth method to screen for enantioselective enzymes without the need to harvest individual bacterial colonies.[56] The *ee*-assay was based on the esterase-catalyzed enantioselective hydrolysis of a fluoroacetate of

pactolactone which leads to fluoroacetic acid – a toxic compound which inhibits the growth of esterase producing yeast. Cell density measurements (OD-values) of the yeast in culture medium were used to monitor the growth rate of the yeast. As yeast was found to be more selective towards the (*S*)-enantiomer, the culture medium containing the (*S*)-enantiomer showed an OD-value that was significantly lower than that of the (*R*)-enantiomer. The difference in the OD-values indicates the enantioselectivity of the yeast.

Significant advances have been made on the methods of chiral identification and quantification based on mass spectrometry. In studies involving kinetic resolution of racemates with ^{2}D isotopic labeling, Reetz et al described an *ee*-assay based on ESI-MS which was used to quantify the *ee* of two lipase-catalyzed kinetic resolution examples: the hydrolysis of pseudo-racemic 1-phenylethyl acetate and the stereoselective esterification of a 1:1 pseudo-racemic mixture of 2-pheynylpropionic acid.[57] In both cases, the (*R*)-enantiomer was labeled with the ^{2}D isotope. About 1000 *ee*s determinations can be performed per day at an accuracy of ±5% (when compared to results from GC). The isotope-based MS assay has also been applied in the directed evolution of enantioselective epoxide hydrolase from *Aspergillus niger,* which were catalysts for the hydrolysis of racemic substrate (*R*)-styrene oxide and its deuterated pseudo enantiomer, (*S*)-D_8-styrene oxide.[58] The ratio of the pseudo-enantiomeric products can be measured by ESI-MS.

Another version of the isotope-labeling approach is NMR-based and dependent on the ^{13}C-labeling of the (*S*)-enantiomer to create a pseudo racemic mixture of 1-phenylethyl acetate and 2-pheynylpropionic acid respectively.[59] Upon catalytic hydrolysis, two pseudo enantiomers in the product mixture can be clearly

distinguished by ^{1}H NMR spectroscopy and the *ee* value can be obtained by analyzing the NMR peaks. A throughput of 1400 *ee* determinations per day at an accuracy of ±2-5% is possible with this method.

The general principle of isotope labeling is further applied to a Fourier transformed infrared spectroscopy (FTIR)-based assay.[60] ^{13}C-labeling of the carbonyl groups was selected to allow easy reading and analysis of the vibrational bands in the IR spectrum. Pseudo-enantiomeric substrates used in this example were (*R*)-1-phenylethyl acetate and (*S*)-(1-phenylethyl)-1-^{13}C-acetate, as well as (*R*)-*N*-1-phenylethylacetamide and (*S*)-*N*-(1-phenylethyl)-1-^{13}C-acetamide. The shift of the respective carbonyl stretching vibration allowed the quantification of the pseudo-enantiomers. Lambert-Beer's law was applied in calculating the concentrations of these pseudo-enantiomers, thus requiring the determination of the molar coefficients of absorbance. This FTIR method is advantageous over the previously described NMR and MS approaches as the *ee* values in culture supernatants could be determined directly. Using an automated system, up to 10000 samples per day is possible by measuring the *ee* values in culture supernatants with an accuracy of ±7%. This method can be extended to the catalytic desymmetrization of prochiral compounds containing isotope labeling.

Assays for enantioselective biotransformation from prochiral substrate

High-throughput assays have also been developed for screening enantioselective synthesis by which the catalyst forms a chiral product from an achiral substrate.

Single stage mass spectrometry methods which include isotope labeling in studies involving desymmetrization of prochiral compounds bearing reactive enantiotopic groups have been introduced. A typical example is provided by a lipase-catalyzed enantioselective hydrolysis of pseudo-meso-1,4-diacetoxy-2-cyclopentene (labeled with ^{2}D) which afforded a mixture of pseudo-enantiomeric products that can be analyzed by MS.[57] Although this method would involve the pre-preparation of pseudo-prochiral compounds, once this task is done, a large number of asymmetric reactions by a huge library of enzymes can be conducted in a high-throughput manner. In an industrial application, Diversa used this MS-based approach in the nitrilase-catalyzed desymmetrization of a prochiral nitrile (3-hydroxyglutaryl nitrile).[61] One of the nitrile functional groups of the substrate was labeled with ^{15}N and the substrate was hydrolyzed to its respective pseudo-enantiomeric products.

The NMR-based approach can also be used to analyze the *ee* of the lipase-catalyzed enantioselective hydrolysis of pseudo-meso-1,4-diacetoxy-2-cyclopentene example.[59] This time, ^{13}C isotope is used to label the prochiral compound. High-throughput NMR analysis with up to 1400 *ee* determinations per day can be achieved via miniaturization and automation of the samples.

There are not many assays available for screening enantioselective reactions for prochiral substrates as it is easier to develop an assay to evaluate the different enantiomers in the product rather than the achiral substrate.

Assays to quantify the enantiomeric products from an enzymatic reaction

A general method for screening enantioselective syntheses is to analyze the *ee* of the enantiomeric products. This method is independent of the nature of the starting substrate used in the enzymatic reaction.

In a method known as EMDee (enzymatic method for determining enantiomeric excess),[62] Abato and Seto was able to determine the enantiopurity of a library of chiral secondary alcohols by using an enzyme to selectively process one enantiomer of the product from a catalytic reaction. In the paper, diethylzinc was added to benzaldehyde to yield 1-phenylpropanol. The (*S*)-aromatic alcohol dehydrogenase from *Thermoanaerobium sp*. was then used to selectively oxidize the (*S*)-enantiomer of the alcohol. The rate of this process can be monitored by observing the formation of NADPH (which relates to the quantity of (*S*)-enantiomer present in the mixture) by UV/Vis spectroscopy at 340mm. Similarly, (*R*)-aromatic alcohol dehydrogenase from *Lactobacillus kefir* can be used to quantify the amount of (*R*)-enantiomer present. About 100 samples can be processed in 30mins thus leading to about 4800 *ee* determinations per day with an accuracy of $\pm$10%. The generality of the EMDee assay in this case may be limited to chiral alcohols which show high enantioselectivity in the oxidation catalyzed by alcohol dehydrogenase. Otherwise, one would have to look for a different, more selective alcohol dehydrogenase. The assay would, of course, need to be optimized for every different chiral alcohol and its related enzyme.

Enzyme immunoassays (EIA), which is widely used in the area of biology and diagnosis, can also be applied in the field of chiral chemistry. Mioskowski and

coworkers demonstrated this new method using a 96-well microtiter plate in which benzoyl formic acid (BF) was reduced in the presence of organometallic catalysts into chiral mandelic acid (MA).[63] After 14 hours, the reaction mixture was then transferred to two different 96-well plates: Plate 1 contained an immobilized antibody that binds both enantiomers of the product and can be used to determine yield; whereas plate 2 contained an immobilized enantiospecific antibody (with high affinity for (*S*)-MA) that can be used to determine enantioselectivity. Both reaction yield and *ee* were measured by determining the decrease in UV/Vis absorbance in their respective plates. The application of this immunoassay method depends on the availability of specific antibodies, which generally can be raised to almost any compound of interest. About 1000 *ee* determinations are possible per day with accuracy of about ±9%.

Belder, Ludwig, Wang and Reetz recently introduced a novel integrated approach which combines a microfluidic reactor and the microchip electrophoresis on a single chip.[64] The hydrolytic kinetic resolution of chiral glycidyl phenyl ether, utilizing different epoxide hydrolase mutants from *Aspergillus niger*, was investigated as a model system. Borate buffer containing heptakis-6-sulfato-β-cyclodextrin as a chiral selector was utilized in the separation of the compounds into respective enantiomers in less than 90s. The chip prototype contains four microvials within the reaction part of the device, thus allowing the reaction and screening of up to three catalysts versus one substrate in a single set of experiment. Detection of educts and products was realized by deep UV native fluorescence detection. Furthermore, integrated chip electrophoresis allowed *ee* value determinations with relative standard deviations (RSD) of 1-3% and E values RSD of 3-21%.

The examples covered in this literature review, though not exhaustive, illustrate a diversity of high-throughput assays that can be used to evaluate the enantioselectivity of enzymes. No single assay is truly universal and at times, high-throughput is obtained by sacrificing accuracy. As the demand for enantiopure compounds continue to grow, the pressure to discover new bio-catalysts via directed evolution and rational design coupled with high-throughput screening/selection methods will push the advances in this field to greater heights. Progress in high-throughput screening/selection will be measured not only in accuracy or applicability, but more importantly in how fast and simple the assay really is.

1.5 Cofactor Regeneration

Most of enzymes which are currently applied in industry are hydrolytic in nature, which means that they perform relatively straight-forward cofactor-independent chemistry.[65-67] In contrast, cofactor-dependent enzymes, such as oxidoreductases and transferases, can perform more complex chemistry and play an important role in many synthetic applications.[67-72] For example, NAD-dependent oxidoreductases catalyze the asymmetric reduction of carbonyl groups to alcohols and amines[72,73] and coenzyme A-dependent (CoA) synthetases catalyze asymmetric carbon-carbon bond formation.[74-77] However, as these enzymes use expensive cofactors, they are rarely applied in industrial settings.

Low molecular weight cofactors like NAD and CoA are essential for numerous enzymatic reactions. Some cofactors such as pyridine dinucleotides (NAD(P)(H)), acetyl coenzyme A, and nucleoside triphosphates (NTPs) act more like cosubstrates and are loosely bound (K_D values in the μM-mM range). They operate as

functional group transfer agents , and are consumed in stoichiometric amounts.[78] Other cofactors which are tightly bound to the enzymes such as adenosylcobalamin, pyridoxal phosphate, biotin, and flavins, are mostly self-regenerating.

The use of enzyme-cofactor reactions is limited by the high costs involved in the stoichiometric addition of cofactors, hence necessitating the need for an efficient *in situ* cofactor regeneration system. Various methodologies of cofactor regeneration have been developed including biological, enzymatical, electrochemical, chemical and photochemical methods.[79-87] An example of a chemical method is in the regeneration of acetyl-CoA from CoA and (*S*)-acetylthiocholine iodide, in which a TTN of 1160 was obtained in the production of citrate.[88]

Enzymatic regenerative strategies are particularly preferred for industrial processes as they have several distinct advantages over their electrochemical, chemical, and photochemical counterparts that include high selectivity, compatibility with production enzymes, better stability, higher TTNs, and superior productivity.[78,86,87] Enzymatic methods for cofactor regeneration have been described for oxidoreductions with NAD(P)(H),[80,89-91] phosphoryl transfer reactions with nucleoside di- and triphosphates (NDP, NTP),[92-97] glycosylations with sugar nucleotides,[98-100] sulfuryl transfer reactions with PAPS,[101,102] and acyl transfer reactions with acetyl-CoA.[74,77,88] Many of these have been successfully implemented in large-scale synthesis.[84,103-105]

1.5.1 NAD(P)H Regeneration

The reduced nicotinamide cofactors NADH and NADPH are involved in many biochemical oxidation and reduction reactions. Unfortunately, the high cost of these

cofactors and their required stoichiometric addition to a reaction is the major inhibitory factor for practical applications . Thus, an *in situ* regeneration method is necessary and its advantages include prevention of product inhibition by the cofactor, simplified reaction workup, and in some cases, a favorable influence on the reaction equilibrium. Enzymatic methods are presently the preferred method for NAD(P)H regeneration, with the current state of the art being the formate/formate dehydrogenase system[106] (FDH) as used on the preparative scale by Degussa for production of L-*tert*-leucine.[107] The advantages of FDH regeneration system include the use of formate as an inexpensive, stable, innocuous substrate and the production of CO_2 that can be easily removed from the reaction. However, disadvantages of this system is its low specific activity (~6 U/mg)[108] and general sensitivity to organic solvents.

Although NADH regeneration by FDH has been most widely applied, other systems have also been developed for NAD(P)H regeneration. Noteworthy non-enzymatic systems for NAD(P)H regeneration typically involve ruthenium or rhodium catalysts which transfer reducing equivalents to $NAD(P)^+$.[109-116] However, these methods generally suffer from catalyst or electrode fouling, or low productivity.[111,112,114,116,117] Other enzymatic methods that have been developed include glucose/glucose dehydrogenase[118] (GDH), glucose/glucose-6-phosphate dehydrogenase[119] (G6PDH) and isopropanol/*Pseudomonas* alcohol dehydrogenases (ADH) for NAD(P)H regeneration.

The recently discovered phosphite dehydrogenase (PTDH) from *Pseudomonas stutzeri*[120] may have kinetic and practical advantages for NADH regeneration. This enzyme catalyzes the oxidation of phosphite to phosphate. The large change in free

energy of this reaction ($\Delta G°'$ = -63.3 kJ/mol estimated from redox potentials) and the associated high equilibrium constant (K_{eq} = 1 x 10^{11}) makes PTDH a promising NADH regenerative enzyme.[121] PTDH also displays a broad pH-rate maximum and the enzyme is not inhibited by phosphate, even at high concentrations.[120] Although PTDH naturally prefers NAD^+ over NADP by about 100-fold, a mutant PTDH has been engineered with relaxed specificity toward both nicotamide cofactors.[122] In a recent work, a thermostable and highly active PTDH was also engineered for practical NAD(P)H cofactor regeneration.[33,123] PTDH was also applied for NADH regeneration in the production of L-lactate from pyruvate by L-Lactate dehydrogenase, D-lactate from pyruvate by D-lactate dehydrogenase, malate from oxaloacetate by malate dehydrogenase, and trifluoromethanol from trifluoroacetaldehyde by an alcohol dehydrogenase.[121]

1.6 Project Overview

This thesis focuses on a novel P450pyr monooxygenase from *Sphingomonas* sp. HXN-200 which belongs to the CYP153 family. CYP153 enzymes are class I P450 proteins, which requires the presence of an electron-delivering protein system (ferredoxin and ferredoxin reductase protein) to effect their monooxygenase activity (Figure 1.3).

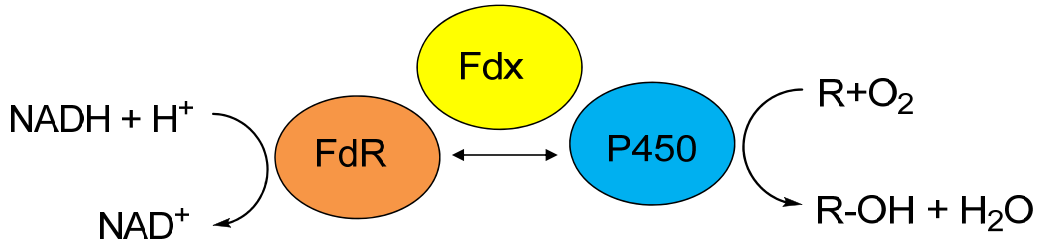


Figure 1.3. Schematic organization of Class I P450s. The ferredoxin reductase (FdR) transfers electrons from NAD(P)H, via an electron shuttle ferredoxin (Fdx), to the heme center of the P450 domain.[11]

This novel enzyme was discovered while screening 70 alkane-degrading strains on a microtitre plate for the 3-hydroxylation of *N*-benzyl pyrrolidine.[124-126] A total of 12 alkane-degrading strains were found to be able to catalyze the hydroxylation of *N*-benzyl pyrrolidine to *N*-benzyl-3-hydroxypyrrolidine. While the well-known membrane-bound AlkB-containing *Pseudomonas oleovorans* GPO1 affords the (*R*)-product in 52% *ee*, *Sphingomonas* sp. HXN-200 exhibits a much higher activity, excellent regioselectivity and the opposite enantioselectivity, giving the (*S*)-product in 53% *ee*. The *Sphingomonas* sp. HXN-200 was also found to be able to catalyze the highly active, regio- and stereoselective hydroxylation of a broad range of substrates, including piperidines, azetidines, 2-pyrrolidinones and 2-piperidinones.[127-130] The enantiomeric products of the hydroxylation reactions are useful pharmaceutical intermediates which are difficult to synthesize via organic chemistry routes. The enzyme responsible, the P450pyr monooxygenase, was thus isolated from this strain.[131]

The enantioselectivity of the wild type (WT) P450pyr for the biohydroxylation of *N*-benzyl pyrrolidine to *N*-benzyl-3-hydroxypyrrolidine is unsatisfactory and therefore it is desirable to improve the P450pyr enantioselectivity via protein engineering. To do so, it would be necessary to first develop a high-throughput *ee*

screening assay which can be applied for asymmetric biohydroxylations. Next, we are also interested in engineering an efficient P450pyr system coupled with a cofactor regeneration to increase activity and productivity. Finally, we will focus on the partial characterization of cytochrome P450pyr and its enantioselective mutants that were previously engineered by directed evolution.

Chapter 2 describes the development of high-throughput screening methods for asymmetric biohydroxylations catalyzed by the P450pyr enzyme. Two high-throughput methods have been developed: 1) two-enzyme-based colorimetric assay and 2) mass spectrometry-based method using an isotopically-labeled optically active substrate. The colorimetric assay was also adapted to a 96-well plate format making it the first example of a high-throughput enantioselectivity assay developed and used in the directed evolution of a monooxygenase for asymmetric biohydroxylations.

Chapter 3 describes the first example where directed evolution has been applied to improve the biohydroxylation enantioselectivity of a class I P450 monooxygenase using a prochiral substrate. Several mutants exhibiting increased and/or inverted enantioselectivity were identified, with product *ee* of 83% (*R*) and 65% (*S*) for mutants 1AF4A and 11BB12, respectively. The homology modeling of the P450pyr and the development of an active *E. coli* recombinant expressing the P450pyr and its electron transport system have also been described.

Chapter 4 describes an efficient monooxygenase biocatalytic system with cofactor recycling which was developed by co-expressing a glucose dehydrogenase from *Bacillus substilis* or a phosphite dehydrogenase from *Pseudomonas stutzeri* together with the P450pyr system in a recombinant *E. coli*.

In chapter 5, the wild type P450pyr and its mutants were purified and reconstituted with their auxiliary electron transport proteins, ferredoxin and ferredoxin reductase *in vitro*. The changes in product *ee* when using 1AF4 and 1AF4A mutants to catalyze the hydroxylations of a range of different substrates were also investigated.

Chapter 2 : Development of a High-throughput Enantiomeric Excess (*ee*) Screening Assay

2.1 Introduction

The expanding world market for chiral fine chemicals and pharmaceuticals has created a great demand for enantiopure compounds. Among the various ways to prepare enantiopure compounds is the use of asymmetric catalysts, particularly biocatalysts or enzymes which are favored due to their high activity and specificity. This has led to the growth in two major areas: 1) the generation and expression of random genetic libraries in suitable host systems and 2) methods for screening or selecting variants of interest that possess the desired characteristic (e.g. enantioselectivity, activity and stability under different process conditions). Vast gene libraries of more than 10^{12} variants can now be readily prepared via molecular biological methods such as error-prone polymerase chain reaction (epPCR) and DNA shuffling. While success in identifying a hit variant has increased, the poor availability of practical high-throughput screening techniques for identifying active variants of interest is a bottleneck in these directed evolution approaches.

Many high-throughput methods for determining the enantioselectivity of enzymes have been developed,[132-138] although most of these methods have been focused on screening for the enantioselectivity in kinetic resolutions since this can be done quickly by comparing the different reaction rates of the two enantiomers of the substrate. The determination of catalyst enantioselectivity for asymmetric transformations from a prochiral substrate is dependent on the measurement of product *ee* except in cases of reversible transformations where their reverse reactions

could be analyzed as kinetic resolutions. The most common methods for analyzing the product *ee* of asymmetric transformations are GC or HPLC with a chiral column which suffers from drawbacks such as a long analytic time and the required work-up procedure to extract the product from the aqueous reaction system. Other medium to high-throughput enantioselectivity analysis method include the use of GC-GC with a chiral column,[139] HPLC with CD/UV or OR/RIU detection,[140,141] chirally modified capillary electrophoresis,[142] electrospray ionization tandem mass spectrometry (MS),[143] color indicators based on doped liquid crystals,[144] or competitive enzyme immunoassays.[145] More complicated high-throughput methods that require the product to be further converted have also been developed. For example, product *ee* can be estimated by exploiting the kinetic resolution effects on the product[146] using mass- or fluorescence-tagged quasi-enantiomeric mixtures of acylating agents with MS[147,148] or fluorescence[149] detection. When the product concentration is known, the product *ee* can be determined by using an enzyme to catalyze a further transformation of the product which can be detected by UV spectroscopy[62] or infrared thermography.[150] The product *ee* can be also established by using two enantioselective enzymes to modify the product with UV detection of NAD(P)H formation.[151] Nevertheless, the application of these methods in catalyst discovery is limited by many factors and can only be applied for certain cases only.

In order to screen the mutant libraries generated in our directed evolution work, we need to develop a practical and efficient high-throughput *ee*-screening assays for asymmetric biohydroxylations. In this chapter we will discuss two high-throughput *ee* screening assays: a two-enzyme-based colorimetric assay[152] and an MS-based method using an isotopically-labeled optically active substrate.[153]

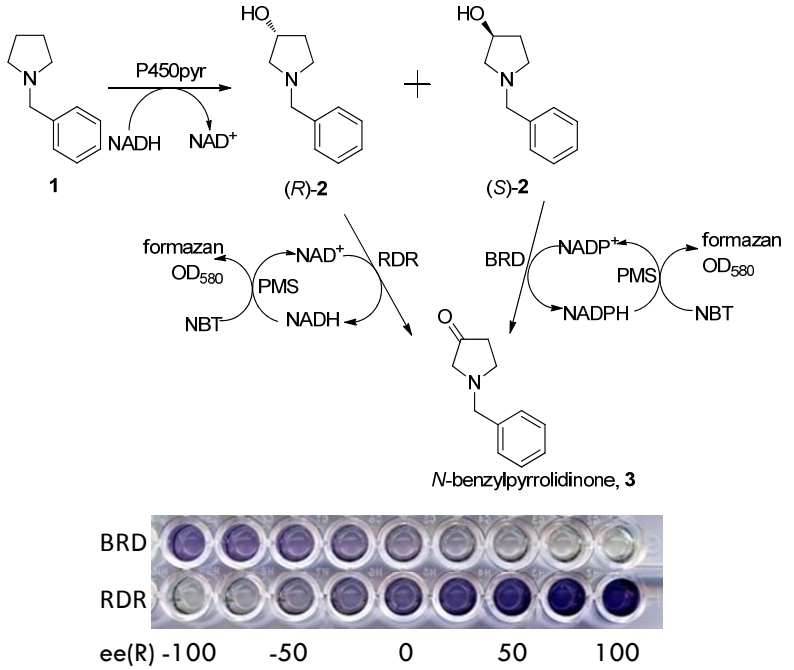
2.2 Two-Enzyme-Based Colorimetric *ee* Screening Assay

2.2.1 Results and Discussion

2.2.1.1 Basic Principle of the Assay

The development of an efficient high-throughput *ee*-screening method was no trivial task. Most of the screening/selection methods that are found in literature focused on improving the enantioselectivity in the kinetic resolution of the enzyme, and very few for asymmetric transformations due to the complexities in developing an *ee*-assay. A high-throughput two-enzyme-based colorimetric assay, adapted to a 96-well plate format, was developed in order to semi-quantitatively determine the *ee* of the racemic product from an asymmetric biohydroxylation. The concept of utilizing two complementary alcohol dehydrogenases has been described,[151,154] but the colorimetric determinations make it the first example of a high-throughput enantioselectivity assay developed and used in the directed evolution of a monooxygenase for asymmetric biohydroxylations.

After an extensive literature search, we found two alcohol dehydrogenases, i.e. BRD from *Micrococcus luteus*[155] and RDR from *Devosia riboflavina*[156] that were highly specific for (*S*)- and (*R*)- 1-benzyl-3- pyrrolidinol **2**, respectively. The principle of our method is illustrated in Scheme 2.1. Enantioselective biohydroxylation of *N*-benzyl pyrrolidine **1** to its corresponding products (*R*)- and (*S*)-*N*-benzyl-3-hydroxypyrrolidines **2** was selected as the target reaction. This example represents an enantioselective hydroxylation of a non-activated carbon atom of a symmetric substrate with equal C-H bond at 3- and 4-position. It is also a useful biotransformation to prepare the corresponding products (*S*)- or (*R*)-**2** which are important pharmaceutical intermediates.[130]



Scheme 2.1. A high-throughput two-enzyme based colorimetric *ee* assay for asymmetric biohydroxylation of prochiral substrate *N*-benzyl pyrrolidine **1** to its corresponding products (*R*)- and (*S*)-1-benzyl-3- pyrrolidinol **2**. The formation of formazan corresponded to the activity of the dehydrogenases that in turn correlated to the concentration of each enantiomer in the aqueous solution.

In the presence of $NADP^+$, the enantioselective alcohol dehydrogenase BRD could convert (*S*)-**2** to 1-benzyl-3-pyrrolidinone **3** whereas in the presence of NAD^+, the RDR could convert (*R*)-**2** to **3**. By utilizing the nitro blue tetrazolium (NBT) - phenazine methosulfate (PMS) colorimetric assay,[122] the production of NAD(P)H was monitored by observing the formation of formazan at OD_{580}. The NBT-PMS-based assay was chosen since it offered simplicity, reproducibility, and a high signal which can be measured even in the presence of whole-cells and media.[122,157] In this assay, the NAD(P)H production was coupled to reduction of NBT_{ox} through the chemical mediator PMS. Upon reduction, NBT_{ox} converts from a soluble yellow compound to a partially soluble purple formazan dye (NBT_{red}). The formation of formazan corresponded to the activity of the dehydrogenases that in turn correlated to the concentration of each enantiomer in the aqueous solution. Hence, the *ee* of the

enantiomeric mixture could be estimated from the apparent concentration of each enantiomer. The *ee* value was also found to be independent of the reaction rate of each alcohol dehydrogenase.

2.2.1.2 Cloning and Expression of the Alcohol Dehydrogenases, BRD and RDR

The BRD gene was cloned from the *M. luteus* gDNA. In order to obtain a higher level of expression of the BRD in *E. coli*, the first 41 codons were optimized to lower the GC ratio[155] without altering the amino acid sequence. The modified BRD gene was cloned into the *NdeI* and *BamHI* sites of the pET15b vector. The BRD was then expressed in *E. coli* BL21(DE3) and purified using a Ni-NTA gravity flow column (Figure 2.1). When the BRD was expressed in TB media, the purified protein was inactive. However, when the BRD Exp media[155] was used, we managed to obtain an active protein.

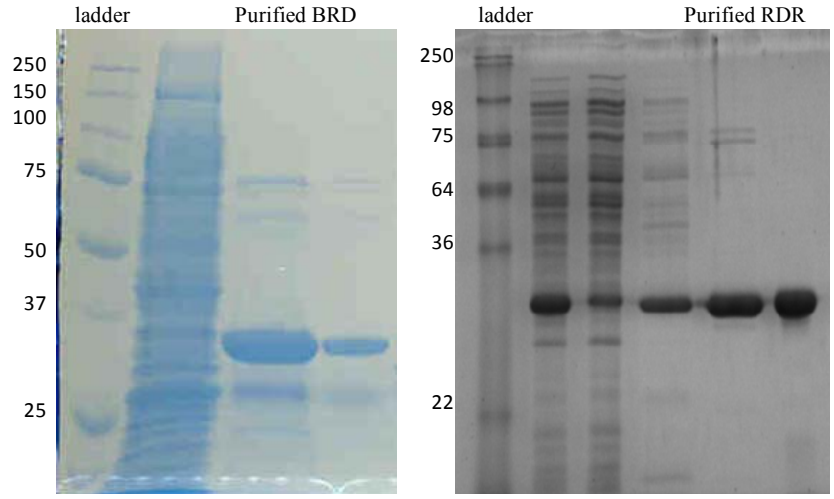


Figure 2.1. SDS-PAGE of purified *N*-histag BRD and *N*-histag RDR.

The RDR could not be amplified successfully by PCR from the *D. riboflavina* gDNA. After several failed attempts, we decided to synthesize and codon optimize the RDR gene by Genscript Corp. (Piscataway, NJ). The sequence of the codon optimized RDR is given in Fig 2.2. The RDR gene was cloned into the *NdeI* and *XhoI* sites of the pET28a vector. As it has been codon optimized, the RDR gene was highly expressed in *E. coli* BL21(DE3) and could be easily purified using a Ni-NTA gravity flow column.

```
  1 ATGAGCCAGG ATTTCAGCGG CAAAGTGGCC TTTGTGACCG GTGGCGCGAG CGGCATTGGT
 61 GAAGCCGTGG TGAAACAGCT GGCGGCGCGT GGTGCGAAAG TTGTGGTTGC GGATCTGAAA
121 CTGGAAGGTG CGCAGGCCGT GGCGGATGCG GTGAAAGCGG CGGGTGGTGA AGCCGCGGCC
181 GTGGCGGTGG ATGTGGCCAA AGCGGATCAG GTGGAAAAAG CGGTTCAGTT TGCGGTGGAT
241 ACCTTTGGTG CGCTGCATCT GGCGGTGAAC AATGCGGGCA TTGGTGGCGC GAGCGCGCCG
301 CTGGGTGATT ATAGCTTTGA TGATTGGCAT CGTGTGATCG ATGTGAACCT GAACAGCGTG
361 TTTTATAGCA TGAAATATGA AATTGTTGCC ATGCTGCGCG CGGGTGGCGG CGCCATTGTG
421 AACATGGCGA GCATTCTGGG TAGCGTGACC TTTCCGAATG CGCCGGCCTA TGTTACCGCG
481 AAACACGGCG TGGTGGGCAT GACCAAAAGC GCCGCCGTTG ATTATGCCAA AAAAGGCATC
541 CGCGTTACCG CGGTTGGCCC GGGCTTTATT GATACCCCGC TGCTGAGCGC CCTGCCGAAA
601 GAAACCCTGG ATTATCTGAA AAGCGTGCAT CCGATTGGTC GCCTGGGCAC CAGCGATGAA
661 GTGGCCGCGC TGACCGCGTT TCTGCTGAGC GATGCGGCCA GCAACATCAC CGGCAGCTAT
721 CATCTGGTGG ATGGTGGTTA TGTGGCCCAG TAA
```

Figure 2.2. Codon optimized sequence of the RDR gene.

2.2.1.3 Assay Evaluation

Initial assay with RDR and BRD showed that they were highly specific to the (*R*)- and (*S*)-**2** respectively in the presence of the relevant cofactors NAD(P)H. They showed no activity in the presence of substrate **1**. To evaluate our assay, we prepared 9 samples containing different mixtures of (*R*)- and (*S*)-**2.** We assayed each sample with purified BRD and RDR separately, coupled with the NBT-PMS assay and found that the *ee* of the sample correlated to the color intensity of the assay (Figure 2.3).

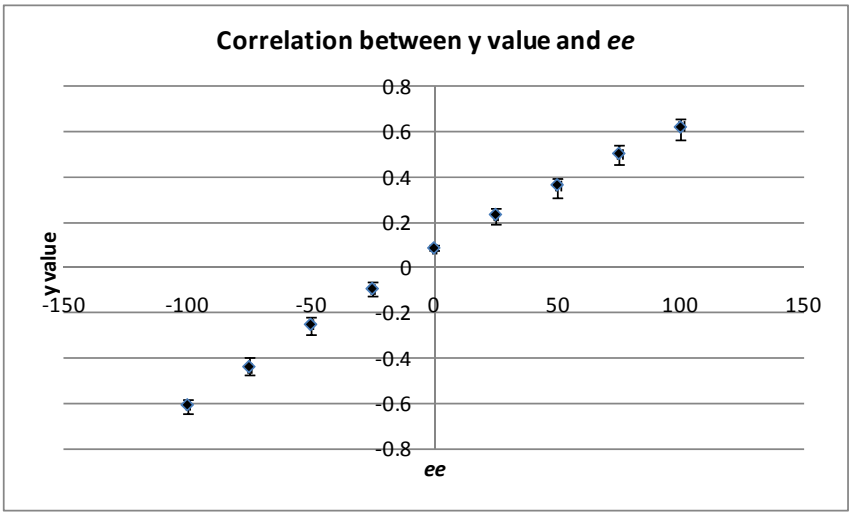


Figure 2.3. Graph shows the linear correlation between y value and *ee* where y where $y = \frac{OD_{580}^{BRD} - OD_{580}^{RDR}}{OD_{580}^{BRD} + OD_{580}^{RDR}}$. $OD_{580}{}^{BRD}$ is linearly correlated to the (*S*)-**2** concentration whereas $OD_{580}{}^{RDR}$ is linearly correlated to the (*R*)-**2** concentration. This experiment was carried out in triplicate.

This assay was further optimized and adapted to a 96-well plate format and checked for consistency. Recombinant *E. coli* BL21(DE3) cells, expressing wildtype P450pyr, ferredoxin and ferredoxin reductase, were cultured and induced in a 96-deep well microplate format. The cells were then used in the biohydroxylation of substrate **1** to product **2** in the microplate. Next, the deep well plates were centrifuged and

aliquots (80 μL) of the supernatant containing the unreacted **1** and its racemic product **2** were pipetted into two separate 96-well microtiter plates. In each plate, the concentrations of (*R*)- and (*S*)-**2** could be determined using the RDR and BRD enzymatic reactions coupled with the NBT-PMS assay. After 2 h, the color of the formazan formed in each well was found to be very consistent. When measured at OD_{580} with a microplate reader, the standard deviation between wells was found to be ~15-20%.

2.2.2 Conclusion and Outlook

A simple yet high-throughput colorimetric assay based on the use of two enantioselective enzymes has been developed to screen the product *ee* of enantioselective hydroxylations. The color intensity of the assay correlates to the *ee* of the enantiomeric mixture, thus enabling us to quickly evaluate the enantioselectivity of the mixture by using a microplate reader.

This high-throughput *ee* screening assay was applied in the directed evolution of an enantioselective P450pyr monooxygenase which will be described further in Chapter 3. To date, none of the known high-throughput enantioselectivity assays in literature has been applied in such directed evolution experiments. Instead, the screening of enantioselective mutants were based mostly on activity assays[158] or low throughput methods such as GC or HPLC analysis with a chiral column.[159] Thus, this work is the first example of a high-throughput enantioselectivity assay developed and used in the directed evolution of a monooxygenase for asymmetric biotransformations.

The advantages of this method includes the following: a) The *ee* can be determined with satisfactory accuracy, independent of the concentration. b) The

enzymes do not have to be specific to, or highly active towards the alcohol. Thus, it is easy to find appropriate alcohol dehydrogenases for analysis of the *ee* value of a given alcohol from the large number of alcohol dehydrogenases with broad substrate ranges that are now available. c) The analysis method is very sensitive and can be used to determine the *ee* value of a sample with a concentration in the μM range, making it very useful for biocatalyst screening, where product concentration is often low. d) No special instruments are required as the analysis is performed with a microplate reader. e) The microplate reader takes only a few seconds to measure the absorbance of each well in a 96-well microtiter plate, thus allowing a high-throughput analysis of the *ee* values. f) The method can be extended to measure the *ee* values of other types of compounds. Two enzymes of another type may be applied, coupled with detection by UV spectroscopy, MS, or even HPLC or GC for short analysis times.

2.3 Mass Spectrometry-Based High-Throughput *ee* Screening Assay

(This part of the chapter was done in collaboration with Dr. Chen Yongzheng.)

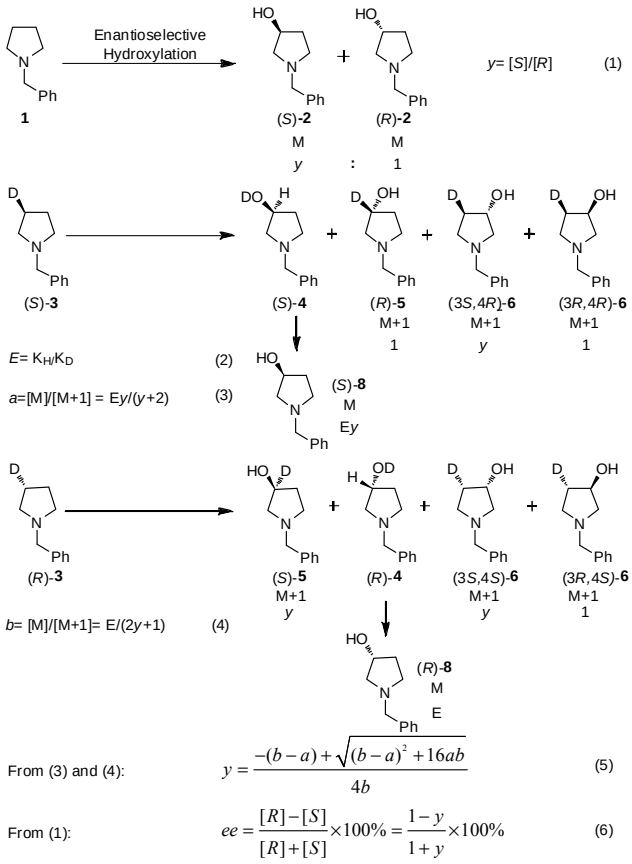
2.3.1 Results and Discussion

MS-based assays are known to be simple and high-throughput, and usually involves the use of isotopically labeled quasienantiomeric or meso substrates[57,160] to determine the catalyst enantioselectivity. However, this approach cannot be extended to the asymmetric transformations of a chiral *non*-meso substrate. Hence, we report a different strategy and new principle for high-throughput *ee* determination for biohydroxylation reactions by the use of optically active, isotopically-labeled substrate with MS detection. The method is novel, and it allows for sensitive, accurate, and high-throughput measurements.

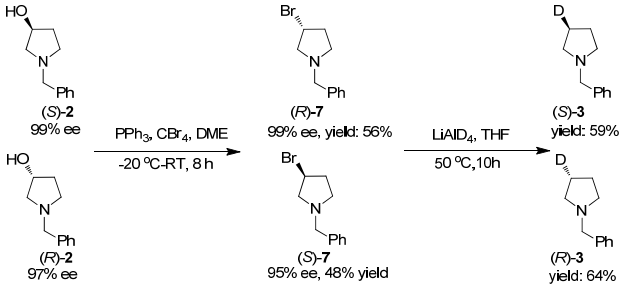
To further demonstrate the generality of this method, enantioselective biohydroxylation of *N*-benzyl pyrrolidine **1** at 3-position was selected as an example. The principle of the method is illustrated in Scheme 2.2. An enantioselective enzyme which could catalyze the hydroxylation of **1** to give the (*S*)-**2** and (*R*)-**2** in a ratio of y is selected. To establish the *y* value and subsequently the product *ee*, a simplified approach is applied. Enantiopure (*R*)- and (*S*)-1-benzyl pyrrolidine-3-d **3** are used as substrates for the biohydroxylation with the same enzyme, respectively. Biohydroxylation of (*S*)-**3** gives (*S*)-**4**, (*R*)-**5**, (3*S*, 4*R*)-**6**, and (3*R*, 4*R*)-**6**, while biohydroxylation of (*R*)-**3** gives (*S*)-**5**, (*R*)-**4**, (3*S*, 4*S*)-**6**, and (3*R*, 4*S*)-**6**. The ratio of these compounds in the products depends on the enzyme enantioselectivity and the deuterium effect (E) on hydroxylation. (*R*)- and (*S*)-**4** have an O-D bond which is quickly exchanged to O-H in the biotransformation medium, giving rise to (*R*)- and (*S*)-**2**, respectively, with a mass of M. All other products have a mass of M+1**.** The ratio of the intensities of M and M+1 peaks in MS can be determined, which corresponds to the *a* value in equation (3) and the *b* value in equation (4), respectively. Based on equation (5) and (6), the *y* and *ee* values can be established.

To prove this concept, (*S*)- and (*R*)-**3** were prepared for the first time from the commercially available enantiopure (*S*)- and (*R*)-**2** according to the routes shown in Scheme 2.3. S_N2 substitution of (*S*)- and (*R*)-**2** using PPh_3/CBr_4 gave (*R*)-**7** in 99% *ee* and 56% yield and (*S*)-**7** in 95% *ee* and 48% yield, respectively. Deuteration of (*R*)- and (*S*)-**7** *via* S_N2 substitution with lithium aluminum deuteride afforded (*S*)-**3** in 64% and (*R*)-**3** in 59% yield, respectively. The *ee* of **3** is expected to be the same as that of the corresponding precursor **7**: (*S*)-**3** in 99% *ee* and (*R*)-**3** in 95% *ee*.

To evaluate the new method, (*S*)-**3** in 99% *ee* and (*R*)-**3** in 95% *ee* were used, respectively, for the hydroxylation. Instead of GC-MS, the potential of using LC-MS for the analysis was explored. At first, *Sphingomonas* sp. HXN-200, a known biocatalyst for the hydroxylation of *non*-activated carbon atom,[130] was selected as the catalyst. Biohydroxylation was performed with 5 mM (*R*)- and (*S*)-**3**, respectively, in a cell suspension of 5.5 g cdw/L in 5 mL 50 mM potassium phosphate buffer (pH 7.5) containing 2% (w/v) glucose at 30 °C for 1 h. The samples directly taken from aqueous biotransformation mixtures were analyzed by LC-MS, and typical LC-MS chromatograms are shown in Figure 2.4. No column separation was required, and the *a* and *b* values were easily obtained from the intensities of the signals at 178 (M) and 179 (M+1). Thus, the *y* and *ee* values were calculated based on equation (5) and (6). As summarized in Table 2.1, the product *ee* for biohydroxylation of **1** with *Sphingomonas* sp. HXN-200 was estimated at 53% (*S*) with this method (Entry 1). The real product *ee* for the biohydroxylation of **1** with *Sphingomonas* sp. HXN-200 under the same conditions was established as 54% (*S*) by chiral HPLC analysis. This demonstrates the high accuracy of the new method. Moreover, the new method using LC-MS analysis does not require the extraction of the product into organic solvent, allowing the direct use of aqueous samples taken from biotransformation mixtures. It also has high sensitivity, allowing the use of a 20× diluted sample containing 0.023 mM product. In comparison, chiral HPLC analysis required the extraction of the product from aqueous medium with ethyl acetate followed by concentration of the sample via evaporation to 30 μL. In fact, chiral HPLC analysis failed with a non-concentrated sample due to the sensitivity limitation. LC-MS analysis took only 1.0 min per sample which equates to a theoretical throughput of 1440 samples per day. In comparison, chiral HPLC analysis required about 30 min per sample.



Scheme 2.2. The principle of a high-throughput enantioselectivity assay for the biohydroxylation of a symmetric substrate based on the use of enantiopure deuterated substrates and MS detection.



Scheme 2.3. Synthesis of (*R*)- and (*S*)-1-benzyl pyrrolidine-3-d **3.**

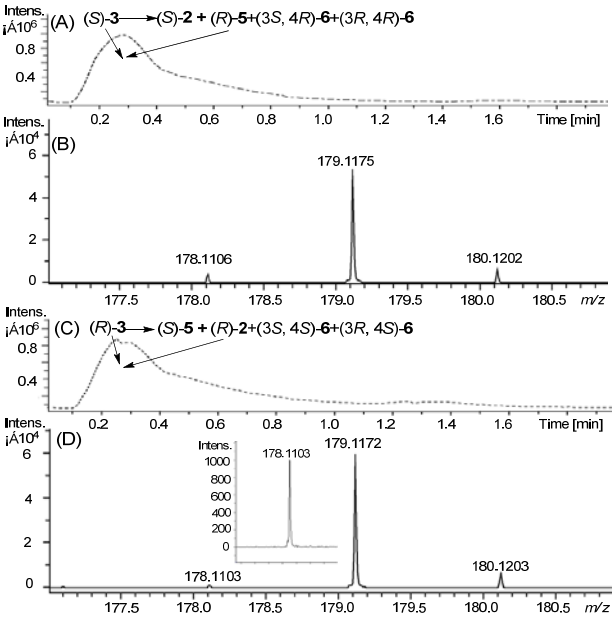


Figure 2.4. LC-MS analysis of the product from biohydroxylation of (*R*)- and (*S*)-**3** with *Sphingomonas* sp. HXN-200, respectively. (A) and (B): Biohydroxylation of (*S*)-**3**; (A) LC chromatogram; (B) MS of the peak between 0.2-0.3 min in LC. (C) and (D): Biohydroxylation of (*R*)-**3**; (C) LC chromatogram; (D) MS of the peak between 0.2-0.3 min in LC.

Table 2.1. Product *ee* of the biohydroxylation of **1** to **2** with different biocatalysts established by an LC-MS-based assay.

Entry	Biocatalysts [a]	Time (h)	Product concentration[b]	a	b	y	*ee* [c] (%)	*ee* [d] (%)
1	*Sphingomonas* sp. HXN-200	1.0	0.023 mM	0.080	0.017	3.25	53(*S*)	54(*S*)
2	*E. coli* BL21(DE3) 1AF4	0.5	0.005 mM	0.049	0.179	0.372	46(*R*)	42(*R*)
3	*P. oleovorans* GPo1	12	0.009 mM	0.014	0.061	0.324	51(*R*)	57(*R*)

[a] Enantioselective hydroxylations of 2-5 mM (*R*)- and (*S*)-**3** were performed, respectively, in 5 mL cell suspension (5-30 g cdw/L) of different biocatalysts at 30°C and 300 rpm. [b] Concentration refers to the analytic sample that has been diluted 20× from the original aqueous sample taken from biotransformation mixtures. [c] Determined and calculated based on LC-MS analysis using the new method. [d] Determined by chiral HPLC analysis of the product from biohydroxylation of **7** with different biocatalysts under same biotransformation conditions. The analytic sample was prepared by extracting the product with 5 mL ethyl acetate and concentrating via evaporation to 30 µL. Without concentrating the sample, it was not possible to determine the product *ee* due to sensitivity limitation.

The method was further examined with other available biocatalysts for this hydroxylation. Recombinant *E. coli* BL21(DE3) expressing 1AF4 mutant of the P450 monooxygenase of *Sphingomonas* sp. HXN-200 catalyzed the biohydroxylation of **1**, giving the corresponding (*R*)-**2** in 42% *ee*. This is the opposite enantioselectivity to *Sphingomonas* sp. HXN-200. To obtain the product *ee* value by our new method, biotransformation of (*R*)- and (*S*)-**3** with the mutant 1AF4 was performed under the same condition, respectively, and the sample from aqueous buffer was analyzed by LC-MS without column separation. As shown in Table 2, Entry 2, the product *ee* was established as 46% (*R*) by this method, which has error of only 4% *ee*. Also in this case, the *ee* determination is sensitive, with product concentration of 0.005 mM, and fast, with an analysis time of 1.0 min.

Pseudomonas oleovorans GPo1 contains a well-known membrane-bound AlkB hydroxylase system which catalyzes the hydroxylation of a number of aliphatic compounds.[161] Biohydroxylation of **1** with this strain gave (*R*)-**2** in 57% *ee*. Similarly, the *ee* value was also established by using the new method via the separate biotransformation of (*R*)- and (*S*)-**3** with *P. oleovorans* GPo1 and LC-MS analysis of the aqueous sample. As shown in Table 2.1, Entry 3, the product *ee* was established as 51% (*R*) by this method with an error of 6% *ee*. Once again, the analysis is sensitive and fast, and also able to analyze the sample containing 0.009 mM product within 1.0 min.

2.3.2 Conclusion

In summary, we have developed a new method for measuring the product *ee* for enantioselective hydroxylation based on the use of enantiopure deuterated

substrates and MS detection. Our method has several distinctive features: a) The *ee* can be determined with satisfactory accuracy, independent of the type of hydroxylations and the nature of biocatalysts. b) The analytic method is very sensitive. It can analyze samples with product concentration as low as 0.005 mM, thus being suitable for the screening of enzymes for biohydroxylation, where the product concentration is often low. c) The analysis is based on MS and does not require separation. It takes only 1.0 min per sample for LC-MS analysis, enabling high-throughput analysis of up to 1440 samples per day. d) The analytic method is simple. It allows for direct analysis of aqueous samples from biotransformations with no further reactions on the bioproduct, a step which is often necessary for many other assays. e) The deuterated substrate can be easily prepared from the corresponding alcohol.

2.4 Materials and Methods

2.4.1 Two-Enzyme-Based Colorimetric *ee* Screening Assay

Materials

Escherichia coli BL21(DE3) and the expression plasmids were purchased from Novagen (Madison, WI). PCR grade dNTPs were obtained from Roche Applied Sciences (Indianapolis, IN). Restriction enzymes, Phusion High-Fidelity DNA polymerase, T4 DNA ligase and their corresponding buffers were purchased from New England Biolabs (NEB) (Beverly, MA). D-Glucose was purchased from ThermoFisher (Pittsburgh, PA). Ampicillin, kanamycin, isopropyl *β*-D-thiogalactopyranoside (IPTG), nitro blue tetrazolium (NBT), phenazine methosulfate (PMS), NAD^+, and $NADP^+$ were purchased from Sigma (St. Louis, MO). Other

required salts and reagents were purchased from either Fisher or Sigma-Aldrich. The Ni-NTA agarose, QIAprep spin plasmid mini-prep kit, QIAEX II gel purification kit, and QIAquick PCR purification kit were purchased from Qiagen (Valencia, CA). Various oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA). SDS-PAGE gels, buffers and protein size markers were purchased from Bio-Rad (Hercules, CA).

Cloning, Overexpression, and the Purification of the BRD Alcohol Dehydrogenase

The genomic DNA from *Micrococcus luteus* was extracted using the Wizard Genomic DNA purification kit from Promega (Madison, WI). The BRD gene was then PCR amplified from the genomic DNA using the following primers: BRDfwd, 5'- TTA ACT ACT CAT ATG CGA CGG ATG ACG CTG CC -3' and BRD rev, 5'-TTA ACT ACT GGA TCC TTA CAG CAT TTC CAG TGG TCG CG -3'. The amplified gene was digested with *Nde*I and *Bam*HI and ligated into the pET15b vector. The first 41 codons were optimized[155] by digesting the pET15b-BRD using *Nde*I and *Xho*I and replacing it with the *Nde*I/*Xho*I digested PCR amplified fragment (5'-tta act CAT ATG CGT CGT ATG ACT TTA CCA TCT GGT GAA TCT ATT CCA GTT TTA GGT CAA GGT ACT TGG GGT TGG GGT GAA GAT CCA GGT CGT CGT GGT GAT GAA GTT GCT GCT TTA CAT GCT GGT CTC GAG tta act-3', *Nde*I and *Xho*I restriction sites are underlined, flanking bases are in small caps). This PCR fragment was synthesized using overlap extension PCR of the following three fragments: fragfwd1 5'-TTA ACT CAT ATG CGT CGT ATG ACT TTA CCA TCT GGT GAA TCT ATT CCA GTT TTA GGT CA-3', fragfwd2 5'-AA GAT CCA GGT CGT CGT GGT GAT GAA GTT GCT GCT TTA CAT GCT GGT CTC GAG TTA

ACT-3', fragrev 5'-ACC ACG ACG ACC TGG ATC TTC ACC CCA CCC CAA GTA CCT TGA CCT AAA ACT GGA ATA GA-3'. The resulting pET15b-BRD* was electroporated into electrocompetent BL21(DE3) and then plated on LB plates containing 100 μg/mL ampicillin. The BL21(DE3) pET15b-BRD* was cultured in 5 mL LB media overnight and then inoculated into 500 mL BRD Exp media[155] (1.5% (w/v) glycerol, 1.5% (w/v) Bacto tryptone, 0.4% (w/v) Bacto yeast extract, 0.2% (w/v) sodium chloride, 0.8% (w/v) potassium dihydrogen phosphate, 0.05% (w/v) magnesium sulfate heptahydrate, pH 6.0). The culture was induced at OD_{600}=0.6-0.8 and cultured at 23°C for 18 h at 250 rpm. The cells were harvested by centrifugation and lysed by French press. The BRD protein was purified using gravity flow columns containing 3 mL of Ni-NTA resin washed with 60 mL 10 mM imidazole buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0), 60 mL 20 mM imidazole buffer, 30 mL 30 mM imidazole buffer and 3 mL 50 mM imidazole buffer. Proteins were eluted with 100 mM imidazole buffer and the protein fractions were analyzed using the Bradford reagent (Bio-Rad). Proteins were concentrated using Millipore Amicon Ultra-4 centrifugal filter device (MWCD 10 kDa) at 4,000 rpm at 4°C, washed three times with 50 mM potassium phosphate buffer (pH 7.5), and concentrated again. Concentrated protein was stored at -80°C in 15% glycerol.

Cloning, Overexpression and the Purification of the RDR Alcohol Dehydrogenase

The RDR gene[156] from *Devosia riboflavina* was codon optimized and synthesized by Genscript Corp (Piscataway, NJ). The gene was PCR amplified further using forward primer RDRfwd NdeI, TTA ACT ACT CAT ATG TCC CAG GAT TTT TCA GGC AAG GTC (*NdeI* restriction site underlined) and reverse primer

RDRrev XhoI, TTA ACT ACT CTC GAG CTA TTG GGC GAC GTA GCC GC (*XhoI* restriction site underlined). The amplified gene was digested with *Nde*I and *Xho*I and ligated into the pET28a vector. The resulting construct was then electroporated into electrocompetent BL21(DE3) and plated on LB plates containing 50 μg/mL kanamycin. The BL21(DE3) pET28a-RDR was cultured in 5 mL LB media overnight and was then inoculated into 500 mL TB media. The subsequent protein expression and purification steps are the same as that for the BRD alcohol dehydrogenase.

Assay Evaluation

To measure the *ee* value, samples containing the two enantiomers (*R*)- and (*S*)-**2** in various ratios were prepared in microplates and separately oxidized with each of two enzymes, namely BRD and RDR. A mixture containing 80 μL aliquots of each sample, 10 μL of the diluted BRD or RDR (~1 μg of enzymes), 0.5 mM of $NAD(P)^+$ as well as 10 μL of nitro blue tetrazolium (NBT) - phenazine methosulfate (PMS) 10× mixture (20 mg NBT and 1 mg PMS) was pipetted into individual deep wells in a microtiter plate. The plates were left in the dark at room temperature for 2 h. The formation of a soluble purple formazan was measured at OD_{580} in a Spectramax 340PC microplate reader (Molecular Devices, Sunnyvale, CA). This experiment was conducted in triplicate. The same protocol was used to analyze the supernatant of a whole-cell assay.

2.4.2 Mass Spectrometry-Based High-Throughput *ee* Screening Assay

Materials

All chemical reagents and solvents were purchased from Acros Organics, Sigma-Aldrich, Fluka, or Merck and used without further purification. *N*-benzyl pyrrolidine **1** (97%, Acros Organics), (*R*)-1-benzylpyrrolidine-3-ol **2** (97% *ee*, Sigma-Aldrich), (*S*)-1-benzylpyrrolidine-3-ol **2** (99% *ee*, Sigma-Aldrich), pyridine (99.9%, Sigma-Aldrich), $POBr_3$ (98.5%, Fluka), THF (99.9%, Sigma-Aldrich), pentane (99.9%, Sigma-Aldrich), CBr_4 (99%, Sigma-Aldrich), PPh_3 (99%, Sigma-Aldrich), CH_2Cl_2 (Fisher Scientific, AR grade), $LiAlD_4$ (98% D, Sigma-Aldrich).

Analytical Methods

The *ee* values of products were determined using a ShimadzuTM Prominence HPLC on a Daicel^{TM} OB-H chiral column (250×4.6 mm, 5 μm) at 25 °C with a flow rate of 1 mL/min and UV detection at 210 nm. Retention time: 17.137 min for (*R*)-**2** and 26.856 min for (*S*)-**2** (OB-H column, 2% isopropanol:98% *n*-hexane); 5.834 min for (*R*)-**7** and 6.489 min for (*S*)-**7** (OB-H column, 2% isopropanol:98% *n*-hexane).

GC-MS spectra were obtained using Agilent GC-MSD system 6890-5973A with GC-MSD HP-5 MS column (analytical conditions: T_c = 160°C). LC-MS spectra were obtained using Brucker UPLC-microTOF Q MS. Analytical conditions: positive mode, 4 bar nebulizing gas, 4.9 kV spray voltage, drying gas flow 8 L/min, drying gas temperature 200 °C.

^{1}H and ^{13}C NMR spectra were recorded on a Brucker-500 (500/75 MHz) spectrometer using $CDCl_3$ as a solvent and TMS as an internal standard.

Optical rotations were measured in specified solution in a 0.1 dm cell at 20 °C with a Perkin-Elmer 341 Polari meter.

Column chromatography was performed with Merck silica gel 60 (200–300 mesh) as column material. TLC analyses were carried out with precoated Silica Gel 60 F_{254} TLC-plate and the compounds were visualized with UV light.

Synthesis of (*R*)- and (*S*)-1-Benzyl pyrrolidine-3-d 3.

Synthesis of (R)-1-benzyl-3-bromopyrrolidine ***7***

CBr_4 (4.96 g, 15 mmol) was added to a stirred solution of (*S*)-1-benzylpyrrolidin-3-ol **2** (1.77 g, 10 mmol) in anhydrous CH_2Cl_2 (20 mL), and the reaction mixture was stirred for 5 min at -20 ^oC. PPh_3 (15 mmol, 3.93 g) was added slowly over a timeframe of 20 min, and the temperature was slowly increased to room temperature (RT). Then, the reaction mixture was continuously stirred at RT for 7h. The solvent was evaporated, and the crude product was purified by chromatography on a silica gel column (R_f = 0.2, *n*-Hexane/CH_2Cl_2/EtOAc of 100:10:5). 1.36 g (*R*)-1-benzyl-3-bromopyrrolidine **7** was obtained (Purity: 99%, HPLC). Yield: 56%. ^{1}H NMR ($CDCl_3$, 500 MHz): δ= 7.238-7.336 (5H, *m*, Ar-*H*), 4.378 (1H, *m*, C*H*Br), 3.656-3.709 (2H, t, CH_2Ph), 3.164-3.199, 2.848-2.880 (2H, t, *J*= 6 Hz, NCH_2CH), 2.664-2.692), 2.698-2.736, 2.527-2.484, (2H, t, *J*= 7.5 Hz, H_a, H_c), 2.707-2.664,2.206-2.250 (2H, m, H_b, H_d); ^{13}C NMR: ($CDCl_3$, 75 MHz): δ= 138.6, 128.7, 128.3, 127.1, 63.7, 59.9, 52.7, 45.7, 36.4, 29.7; $[\alpha]_D^{20}$= -19.6 (*c* = 1.38, CH_2Cl_2); *ee* = 99% (*R*).

Synthesis of (S)-1-benzyl pyrrolidine-3-d ***3***

$LiAlD_4$ (420 mg, 10 mmol) was added to a stirred solution of (*R*)-1-benzyl-3-bromopyrrolidine **7** (1.2 g, 5 mmol) in anhydrous THF (10 mL),. The mixture was stirred for 10 hours at 50 ^oC. Upon completion of reaction, 1 mL of aq NaOH (1 M) was added dropwise, followed by the addition of 3 mL of water. The product was extracted with pentane (2 × 50 mL). The organic phase was separated, dried with anhydrous $MgSO_4$, filtered, and concentrated by evaporation. After flash chromatography on a silica gel column (R_f = 0.2, CH_2Cl_2: MeOH = 100: 2), 0.47 g (*S*)-**3** was isolated as a colorless liquid (Purity: 99%, HPLC). Yield: 59%. ^1H NMR ($CDCl_3$, 500 MHz): δ= 7.223-7.342 (5H, m, Ar-*H*), 3.615 (2H, s, CH_2Ph), 2.507 (4H, t, *J*= 6.5 Hz, NCH_2), 1.746-1.795 (3H, m, BrC*H*CH_2CH_2); ^{13}C NMR: ($CDCl_3$, 75 MHz): δ= 139.35, 128.9, 128.2, 126.8, 60.7, 54.1, 54.0, 23.3, 23.1.

Synthesis of (S)-1-benzyl-3-bromopyrrolidine ***7***

CBr_4 (4.96 g, 15 mmol) was added to a stirred solution of (*R*)-1-benzylpyrrolidin-3-ol **2** (1.77 g, 10 mmol) in anhydrous CH_2Cl_2 (20 mL), and the reaction mixture was stirred for 5 min at -20 ^oC. PPh_3 (15 mmol, 3.93 g) was added slowly over a timeframe of 20 min, and the temperature was slowly increased to RT. Then, the reaction mixture was stirred continuously at RT for 7h. The solvent was evaporated and the crude product was purified by column chromatography on silica gel (R_f = 0.2, *n*-hexane/CH_2Cl_2/EtOAc = 100:10:5). 1.15 g (*S*)-1-benzyl-3-bromopyrrolidine **7** was obtained (Purity: 99%, HPLC). Yield: 48%. ^1H NMR ($CDCl_3$, 500 MHz): δ= 7.247-7.328 (5H, m, Ar-*H*), 4.355-4.402 (1H, m, C*H*Br), 3.658-3.737 (2H, t, CH_2Ph), 3.167-3.201, 2.849-2.881 (2H, t, *J*= 6 Hz, NCH_2CH), 2.738-2.787,

2.485-2.544, (2H, t, *J*= 7.5 Hz, H_a, H_c), 2.711-2.667,2.206-2.251 (2H, m, H_b, H_d); ^{13}C NMR: ($CDCl_3$, 75 MHz): δ= 138.5, 128.7, 128.3, 127.1, 63.6, 59.9, 52.7, 45.7, 36.4, 29.7.; $[\alpha]_d^{20}$= = +18.5 (*c* = 1.35, CH_2Cl_2); *ee* = 95% (*S*).

Synthesis of (R)-1-benzyl pyrrolidine-3-d ***3***

$LiAlD_4$ (420 mg, 10 mmol) was added to a stirred solution of (*S*)-1-benzyl-3-bromopyrrolidine **7** (1.2 g, 5 mmol) in anhydrous THF (10mL). The reaction mixture was stirred for 10 hours at 50 ^oC. Upon completion of reaction, 1 mL of aq NaOH (1 M) was added dropwise, followed by the addition of 3 mL of water. The product was extracted with pentane (2 × 50 mL). The organic phase was separated, dried with anhydrous $MgSO_4$, and concentrated by evaporation. After flash chromatography on a silica gel column (R_f= 0.2, CH_2Cl_2: MeOH = 100: 2), 0.51 g (*R*)-**9** was isolated as a colorless liquid (Purity: 99%, HPLC). Yield: 64%. ^{1}H NMR ($CDCl_3$, 500 MHz): δ= 7.223-7.342 (5H, m, Ar-*H*), 3.614 (2H, s, CH_2Ph), 2.506 (4H, t, *J*=6.5 Hz, NCH_2), 1.745-1.803, (3H, m, BrC*H*CH_2CH_2); ^{13}C NMR: ($CDCl_3$, 75 MHz): δ= 139.4, 128.9, 128.2, 126.7, 60.8, 54.2, 54.1, 23.8, 23.1.

Biotransformation of 1, (*R*)- and (*S*)-3 with Different Biocatalysts.

Biotransformation of ***1****, R-****3****, S-****3*** *with Sphingomonas sp. HXN-200*

The cells of *Sphingomonas* sp. HXN-200 were cultured in E2 medium with *n*-octane as carbon source[130]. Frozen/thawed cells of *Sphingomonas* sp. HXN-200 were resuspended into 5 mL of 50 mM potassium phosphate buffer (pH 7.5) containing 2% (w/v) glucose to a cell density of 5.5 g cdw/L in a 50 mL Erlenmeyer flask. 5 mM **1**, (*R*)- and (*S*)-**3** were added, respectively, into different flasks. The mixtures were

shaken at 30 °C and 250 rpm for 1 h. For hydroxylation of substrate **1**, the cells were removed from the bioconversion mixture by centrifugation, and the supernatant was adjusted to pH 11-12 by addition of KOH followed by extraction of the hydroxylation product with ethyl acetate. The organic phase was separated, dried over anhydrous $MgSO_4$, and removed by evaporation. The bioproduct was re-dissolved in 0.3 mL isopropanol for chiral HPLC analysis. Product *ee* was 54% (*S*). For hydroxylation of (*R*)-**3** and (*S*)-**3**, the cells were first removed by centrifugation. Then, 50 µL of the supernatant was diluted 20 times with methanol and used as sample for LC-MS analysis.

*Biotransformation of **1**, (R)-**3**, (S)-**3** with 1AF4 mutant*

The 1AF4 mutant was streaked on an LB agar plate containing 100 µg/mL ampicillin and 50 µg/mL kanamycin, and was incubated overnight at 37 ºC. A single colony of the 1AF4 mutant was inoculated into 5 mL of LB medium with 100 µg/mL ampicillin and 50 µg/mL kanamycin, and cultured overnight at 37 ºC. 500 µL of the overnight culture was added into 50 mL of TB medium containing 100 µg/mL ampicillin and 50 µg/mL kanamycin, and was shaken at 250 rpm, 37 ºC for 2 h. When OD_{600} reached 0.6-0.8, the culture was induced with 0.5 mM IPTG and shaken at 250 rpm, 30 ºC for another 5 h. The cells were finally harvested at $OD_{600} \sim 4$.

In parallel experiments, the freshly harvested cells were resuspended in 50 mM potassium phosphate buffer (pH 7.5) containing 2% (w/v) glucose to a cell density of 10 g cdw/L in a 100 mL Erlenmeyer flask. 2 mM of **1**, (*R*)- and (*S*)-**3** were added, respectively, into different flasks. The biotransformations were performed at 30 °C, 250 rpm for 12 h. For hydroxylation of substrate **1**, the cells were removed

from the bioconversion mixture by centrifugation, and the supernatant was adjusted to pH 11-12 by addition of KOH followed by extraction with ethyl acetate. The organic phase was separated, dried over $MgSO_4$, and removed by evaporation. The bioproduct was dissolved in 0.3 mL CH_3OH for chiral HPLC analysis. Product *ee* was 42%(*R*). For hydroxylation of (*R*)- and (*S*)-**3**, the cells were first removed by centrifugation. Then, 50 μL of the supernatant was diluted 20 times with methanol and used as sample for LC-MS analysis.

*Biotransformation of **1**, (R)-**3**, (S)-**3** with P. oleovorans GPo1*

P. oleovorans GPo1 was grown on E2-agar plate and then inoculated into 5 mL LB medium in a 28 mL glass bottle with a screw cap. The culture was shaken at 30 ºC and 300 rpm for 7 h and then added to 100 mL E2 liquid medium containing trace elements in a 250-mL shaking flask with ventilated plastic stopper to reach an initial cell density of 0.1 g cdw/L. 15 mL of plastic tube containing 1 mL octane was put into the flask, and the vapor of toluene was used as the carbon source. The culture was incubated at 30 ºC and 300 rpm for 24 h (OD_{450} = 2.7).

In parallel experiments, cells of *P. oleovorans* GPo1 were suspended in 5 mL of 50 mM KH_2PO_4-K_2HPO_4 buffer (pH 7.5) to a density of 30 g cdw/L. 2 mM of **1**, (*R*)-**3** and (*S*)-**3** were added, respectively, to individual cell suspensions. The mixture was shaken at 30 °C and 300 rpm for 0.5 h. For hydroxylation of substrate **1**, the cells were removed from the bioconversion mixture by centrifugation, and the supernatant was adjusted to pH 11-12 by addition of KOH followed by extraction with ethyl acetate. The organic phase was separated, dried over $MgSO_4$, and removed by evaporation. The bioproduct was re-dissolved in 0.3 mL CH_3OH for chiral-HPLC

analysis. Product *ee* was 57% (*R*). For the hydroxylation of (*R*)-**3** and (*S*)-**3**, the cells were first removed by centrifugation. Then, 50 μL of the supernatant was diluted 20 times with methanol and used as sample for LC-MS analysis.

Example of ee determination with LC-MS.

The samples were analyzed by Brucker UPLC-microTOF Q MS (Analytical conditions: positive mode, 4 bar nebulizing gas, 4.9 kV spray voltage, drying gas flow 8 L/min, drying gas temperature 200 °C). As for the product **2** ($C_{11}H_{15}NO$), [M+1]/[M] is 12% due to the existence of ^{13}C.

a) For Entry 1, Table 2.1, the LC-MS spectra were shown in Figure 2.5 and 2.6.

From Figure 2.5: [M] = 1011 [M+1] = 59536

a = [M]/[M+1] = Ey/(y+2) a = [M]/[M+1] = 1011/(59536-12%×1011) = 0.080

From Figure 2.6: [M] = 4233 [M+1] = 53684

b = [M]/[M+1]= E/(2y+1) b = [M]/[M+1] = 4233/(53684-12%×4233) = 0.017

$$y = \frac{-(b-a)+\sqrt{(b-a)^2+16ab}}{4b} = 3.25$$

$$ee = \frac{[R]-[S]}{[R]+[S]} \times 100\% = \frac{1-y}{1+y} \times 100\% = \frac{1-3.25}{1+3.25} = -53\%(R) = 53\%(S)$$

b) For Entry 2, Table 2.1, the LC-MS spectra were shown in Figure 2.7 and 2.8.

From Figure 2.7: [M] = 62 [M+1] = 354

a = [M]/[M+1] = Ey/(y+2) a = [M]/[M+1] = 62/(354-12%×62) = 0.049

From Figure 2.8: [M] = 19.5 [M+1] = 400

b = [M]/[M+1]= E/(2y+1) b = [M]/[M+1] = 4233/(53684-12%×4233) = 0.179

$$y = \frac{-(b-a)+\sqrt{(b-a)^2+16ab}}{4b} = 0.372$$

$$ee = \frac{[R]-[S]}{[R]+[S]} \times 100\% = \frac{1-y}{1+y} \times 100\% = \frac{1-0.372}{1+0.372} = 46\%(R)$$

c) For Entry 3, Table 2.1, the LC-MS spectra were shown in Figure 2.9 and 2.10.

From Figure 2.9: [M] = 62 [M+1] = 1024

a = [M]/[M+1] = Ey/(y+2) a = [M]/[M+1] = 62/(1024-12%×62) = 0.014

From Figure 2.10: [M] = 27.5 [M+1] = 1987

b = [M]/[M+1]= E/(2y+1) b = [M]/[M+1] = 27.5/(1987-12%×27.5) = 0.061

$$y = \frac{-(b-a)+\sqrt{(b-a)^2+16ab}}{4b} = 0.324$$

$$ee = \frac{[R]-[S]}{[R]+[S]} \times 100\% = \frac{1-y}{1+y} \times 100\% = \frac{1-0.324}{1+0.324} \times 100\% = 51\%(R)$$

GC-MS and LC-MS spectra (Figure 2.5-2.10)

Figure 2.5. LC-MS chromatogram of biohydroxylation (*S*)-**3** with *Sphingomonas* sp. HXN-200.

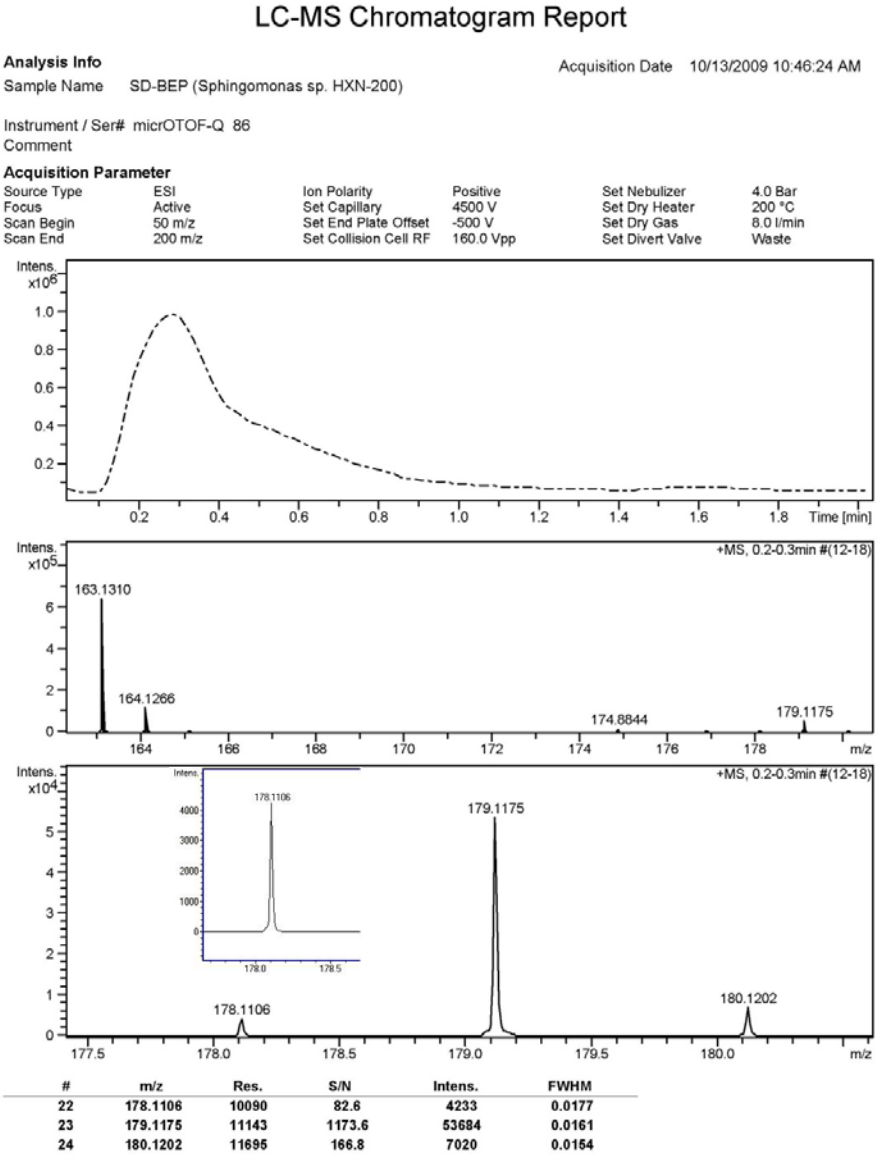


Figure 2.6. LC-MS chromatogram of biohydroxylation (*R*)-**3** with *Sphingomonas* sp. HXN-200.

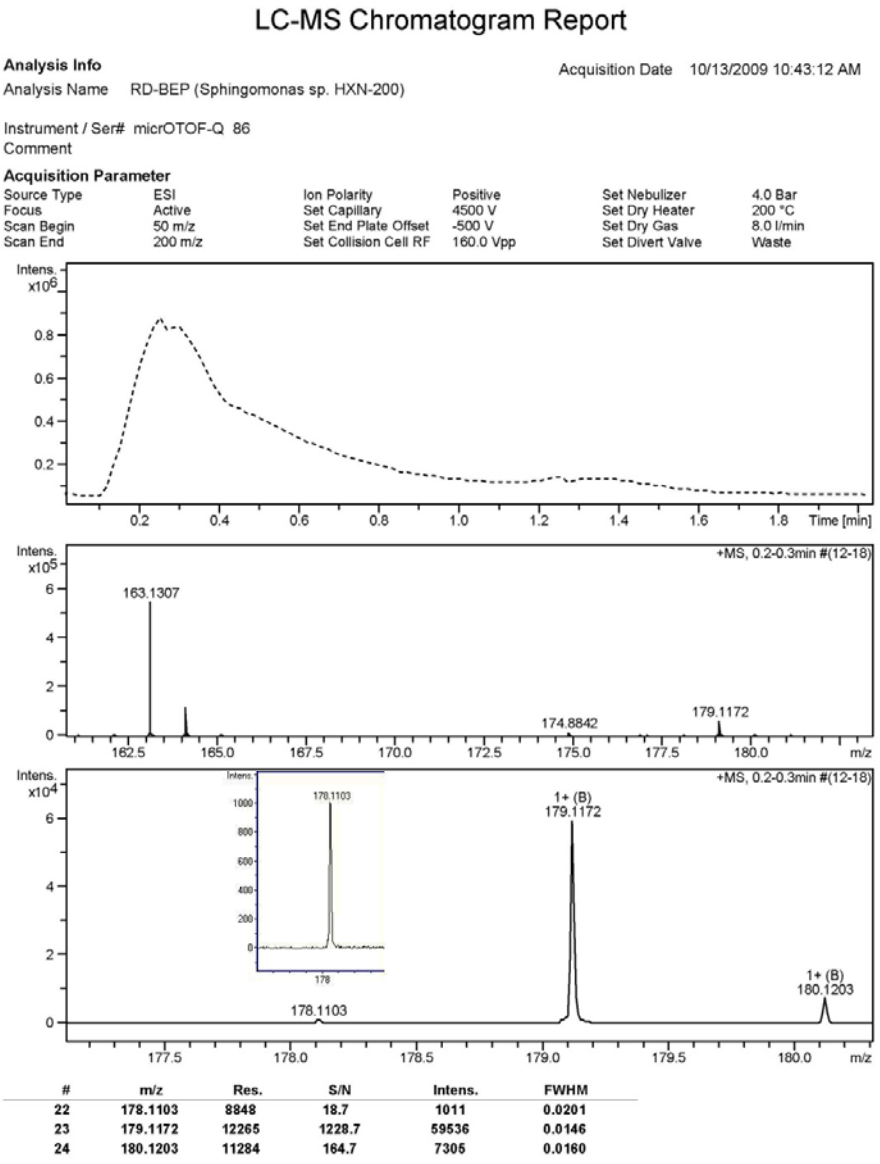


Figure 2.7. LC-MS chromatogram of biohydroxylation (*S*)-**3** with 1AF4

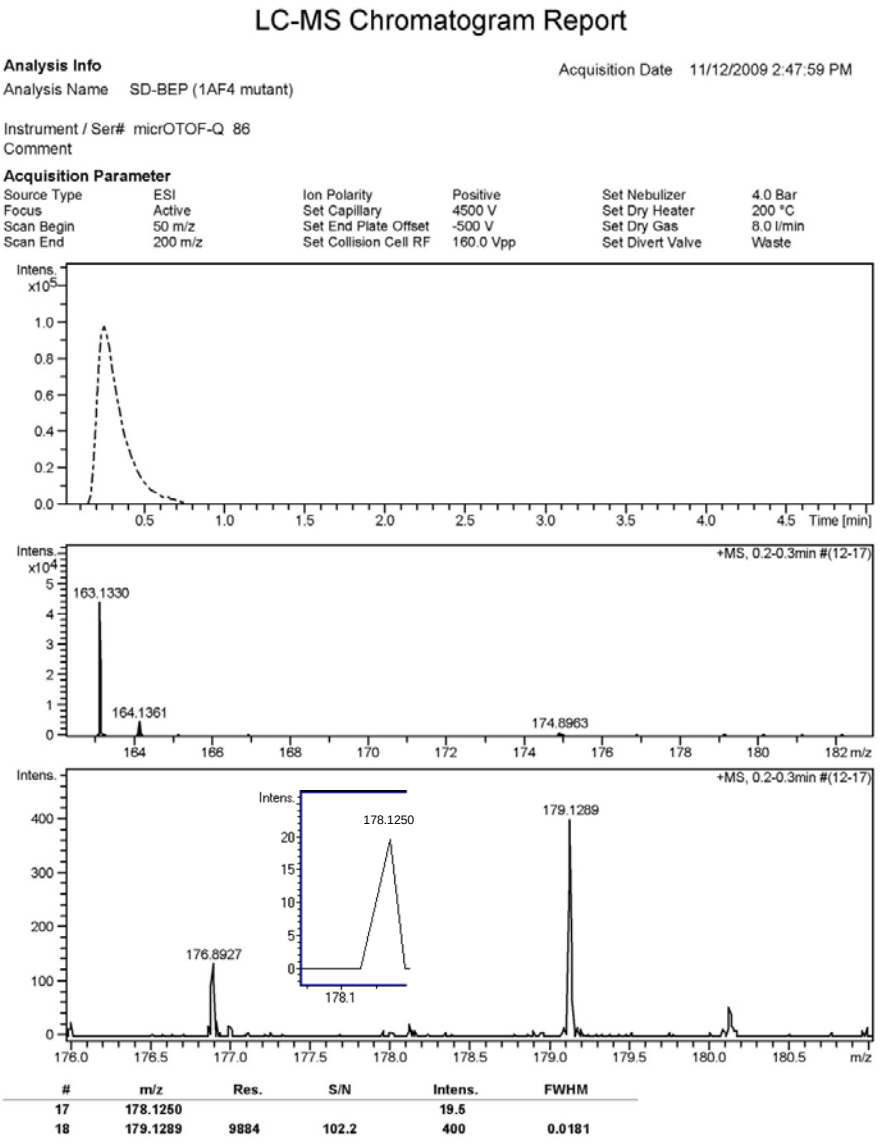


Figure 2.8. LC-MS chromatogram of biohydroxylation (*R*)-**3** with 1AF4

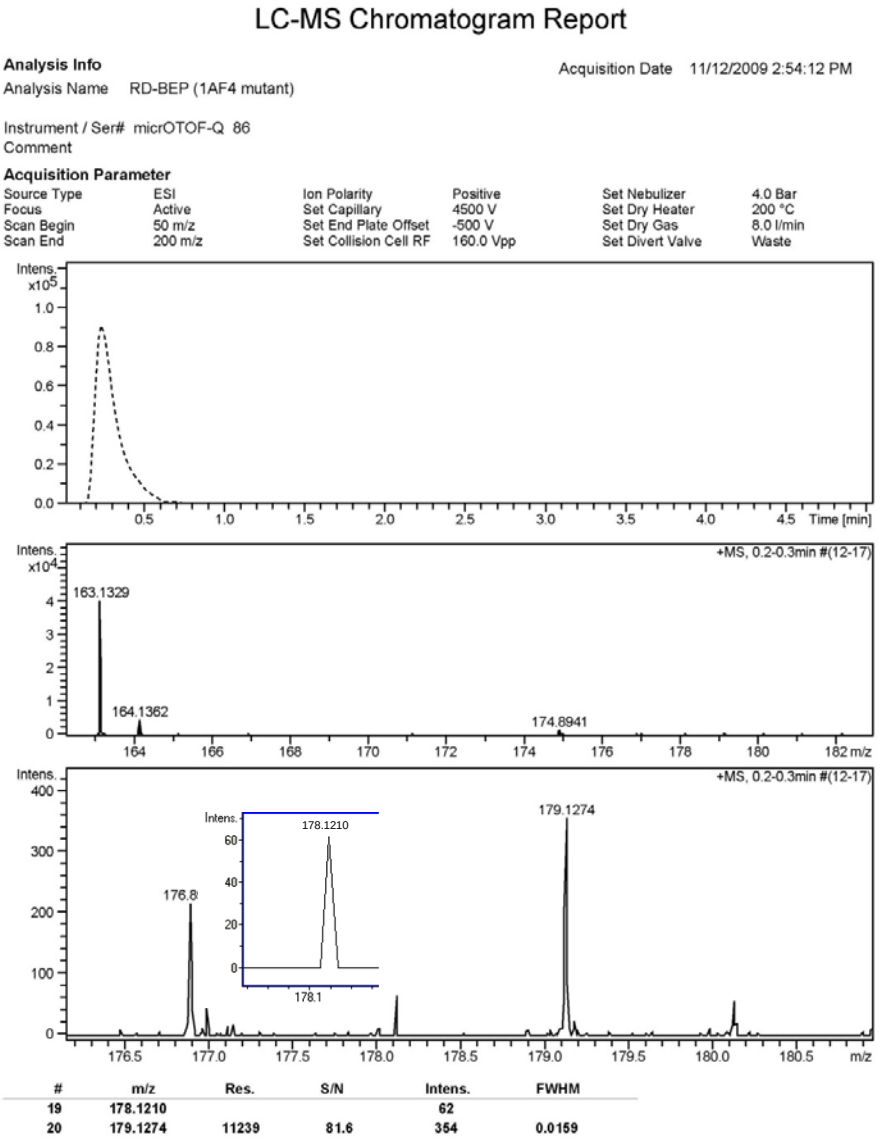


Figure 2.9. LC-MS chromatogram of biohydroxylation (*S*)-**3** with *P. oleovorans* GPo1

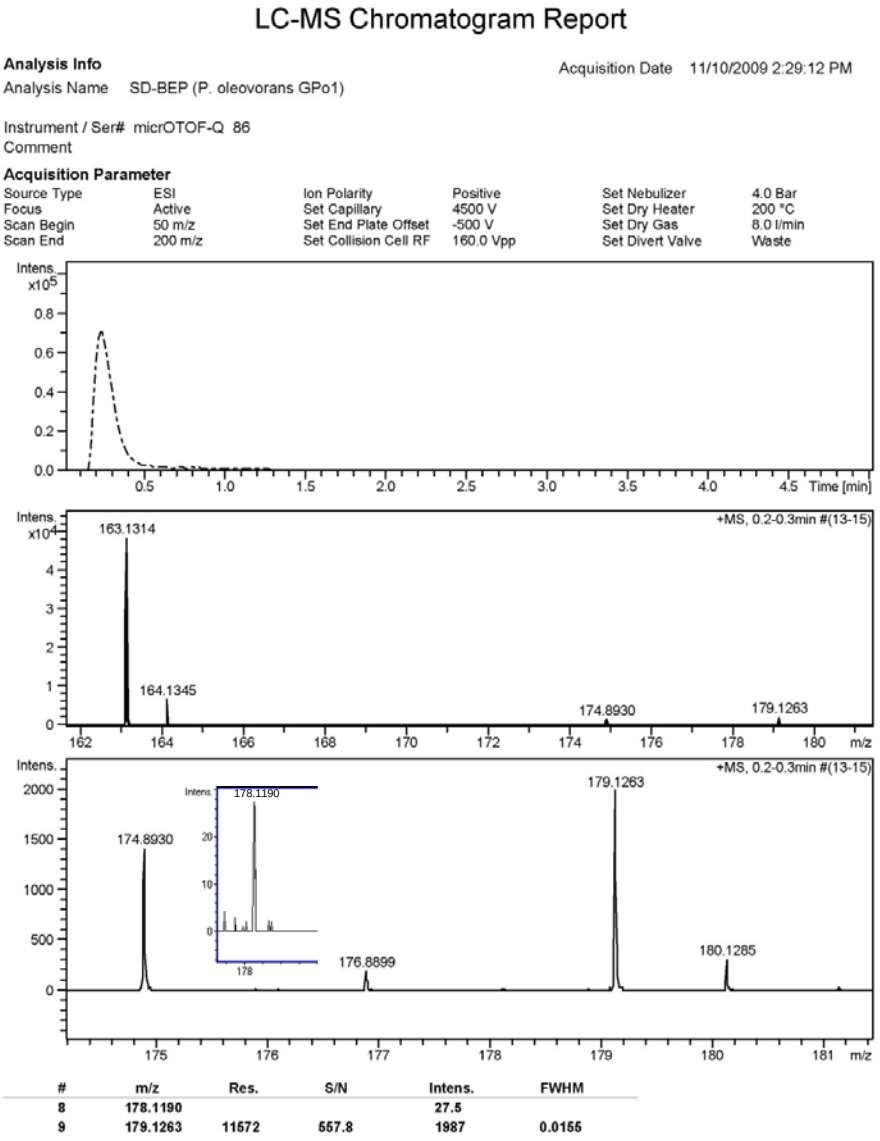
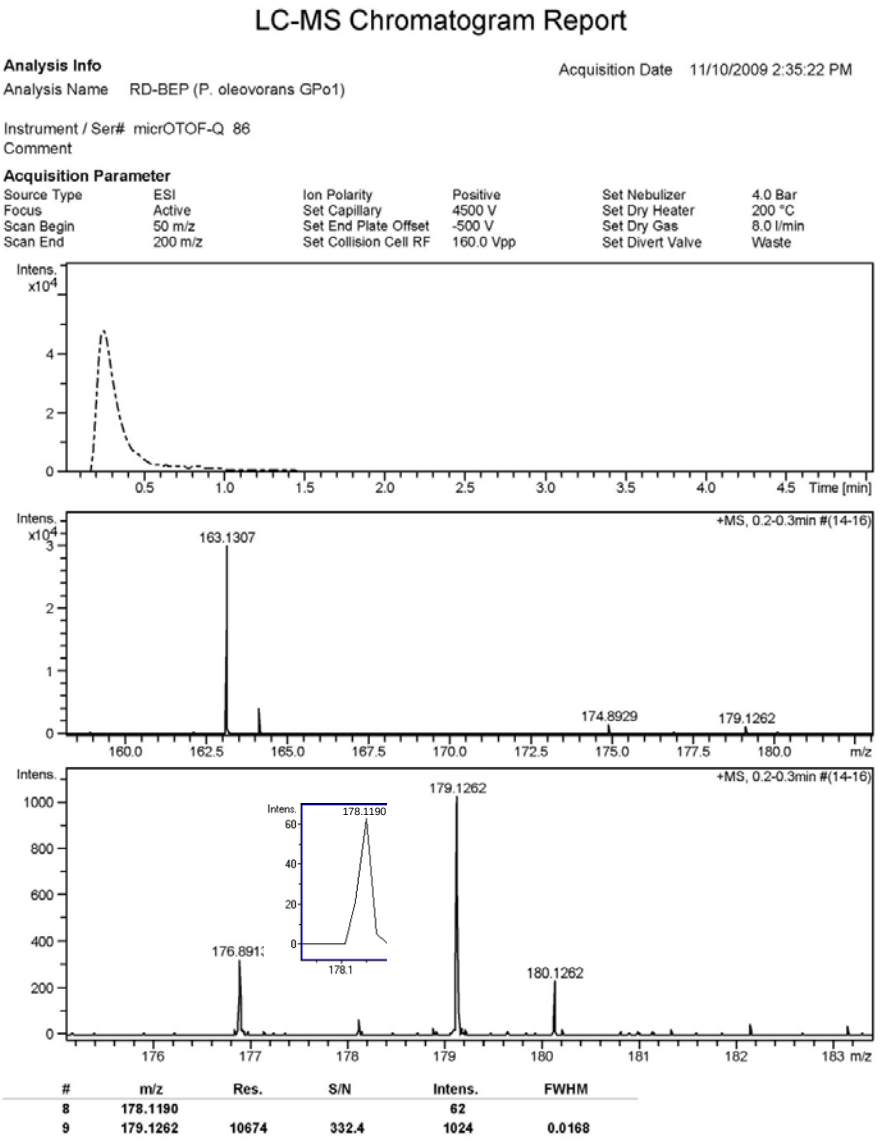


Figure 2.10. LC-MS chromatogram of biohydroxylation (*R*)-**3** with *P. oleovorans* GPo1



Chapter 3 : Inverting the Enantioselectivity of P450pyr Monooxygenase by Directed Evolution

3.1 Introduction

Enantioselective hydroxylation of non-activated carbon atoms represents a significant challenge in classical organic chemistry.[9] However, nature has offered us a solution via biocatalysis with monooxygenases using molecular oxygen as oxidant. Cytochrome P450 monooxygenases (also known as CYP) constitute the largest family of heme-containing monooxygenases that can oxidize a broad range of substrates, often at non-reactive carbon centers.[151] Of particular interest is a novel P450pyr enzyme from *Sphingomonas* sp. HXN-200[126] that belongs to the class I P450 proteins and requires the presence of an electron-delivering protein system (i.e. ferredoxin and ferredoxin reductase) to effect its monooxygenase activity. The P450pyr monooxygenase was found to catalyze the regio- and stereoselective hydroxylation of *non*-activated carbon atom with broad substrate range, high activity, excellent regioselectivity, and good to excellent enantioselectivity. It was able to catalyze the hydroxylations of *N*-substituted pyrrolidines, piperidines, azetidines, 2-pyrrolidinones and 2-piperidinones.[127-130] The hydroxylation enantiomeric products such as (*R*)- and (*S*)-3-hydroxypyrrolidines, (*S*)-4-hydroxy-pyrrolidin-2-ones, 4-hydroxypiperidines, and 2-hydroxyazetidines are useful pharmaceutical intermediates which are difficult to synthesize via organic chemistry routes.

We are particularly interested in the enzymatic hydroxylation of *N*-benzyl pyrrolidine **1** to its corresponding (*R*)- and (*S*)-*N*-benzyl-3-hydroxypyrrolidines **2**.

The (*R*)- and (*S*)-*N*-protected 3-hydroxypyrrolidines are interesting pharmaceutical intermediates. The (*S*)-enantiomers are used in the synthesis of calcium antagonist *Barnidipine* **3**, an agent for irritable bowel syndrome, *Darifenacin* **4**, and anticoagulant DX-9065a **5** (Figure 3.1). The (*R*)-enantiomers are intermediates for the preparation of carbapenem antibiotics RS-533 **6** and antibacterial drug ABT-719 **7**. Also, the racemic 3-hydroxypyrrolidines are useful intermediates in the synthesis of antibacterial drugs *Tosuxacin* **8** and *Clinafloxacin* **9.**[126] Unfortunately, the enantioselectivity of the wild type (WT) P450pyr is unsatisfactory and therefore it is desirable to improve the P450pyr enantioselectivity via protein engineering.

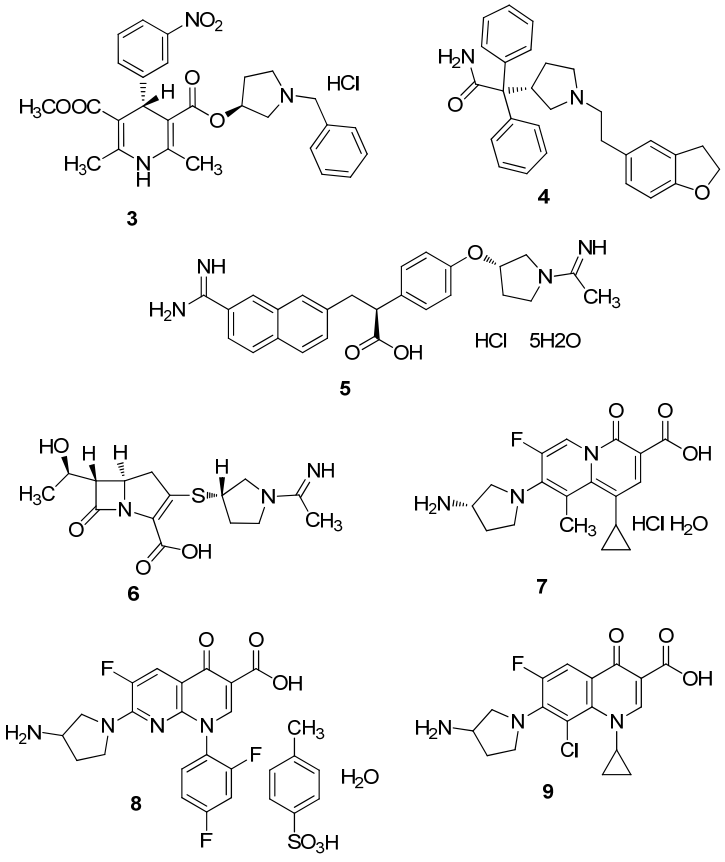


Figure 3.1. Applications of (*R*)- and (*S*)-*N*-protected 3-hydroxypyrrolidines.

Directed evolution, has recently emerged as a powerful approach for engineering biocatalysts.[162-165] As discussed in Chapter 1, this *in vitro* method mimics the Darwinian evolution and involves creation of genetic diversity followed by screening/selection for those variants with desired features. Directed evolution and rational design have been used to enhance the enantioselectivity of enzymes, such as lipases, esterases, hydantoinases, nitrilases, epoxide hydroxylases, phosphotriesterases, aminotransferases, aldolases, cyclohexanone and cyclopentanone monooxygenases, and monoamine oxidases.[159] Previous directed evolution studies on P450 monooxygenases have focused on altering their catalytic efficiency,[43,166] substrate specificity,[44] regioselectivity,[45] solvent tolerance,[46] and thermostability.[47] There have also been many successful attempts to engineer an enzyme for the reversal in enantioselectivity,[167-172] but none of these studies were done with the P450 monooxygenases and very few on asymmetric catalysis. While screening a library of self-sufficient P450 BM-3 mutants for changes in regioselectivity, Arnold and coworkers have also identified some mutants with altered enantioselectivity based on an activity assay.[158] Here, we report the first example where directed evolution, coupled with a high-throughput enantioselectivity screening system, has been applied to improve the biohydroxylation enantioselectivity of a class I P450 monooxygenase using a prochiral substrate.

3.2 Results

3.2.1 Homology Modeling

A protein sequence BLAST search was performed against the Protein Data Bank (PDB) and seven sequences were chosen from the highest scoring results. They were cytochrome P450terp (1CPT) from *Pseudomonas* sp., CYP119 from *Sulfolobus*

solfactaricus (1F4T), P450nor (1XQD) from *Fusarium oxysporum*, P450st from *Sulfolobus tokodaii* (1UE8), P450epok (1PKF) from *Sorangium cellulosum*, CYP158A2 from *Streptomyces Coelicolor* (1S1F), and P450cam (1K2O) from *Pseudomonas putida* which share 24-29% sequence identity with P450pyr. In Figure 3.2, the sequence alignment of these proteins with P450pyr is displayed with the conserved residue highlighted in blue.

Consensus key
* - single, fully conserved residue
: - conservation of strong groups
. - conservation of weak groups
- no consensus

```
P450pyr             MEHTGQSAAATMPLDSIDVSIPELFYNDSVGEYFKRLRKDDPVHYCADS-AFGPYWSITK
P450st_1UE8         -----------------------------MYDWFKQMRKESPVYYDGK------VWNLFK
CYP119_1F4T         -----------------------------MYDWFSEMRKKDPVYYDGN------IWQVFS
P450epok_1PKF       --MTQEQANQSETKPAFDFKPFAPGYAEDPFPAIERLREATPIFYWDE----GRSWVLTR
CYP158A2_1S1F       ----MTEETISQAVPPVRDWPAVDLPGSDFDPVLTELMREGPVTRISLPNGEGWAWLVTR
P450nor_1XQD        ------------MASGAPSFPFSRASGPEPPAEFAKLRATNPVSQVKLFDG-SLAWLVTK
P450terp_1CPT       ---MDARATIPEHIARTVILPQGYADDEVIYPAFKWLRDEQPLAMAHIEG-YDPMWIATK
P450cam_1K2O        --MTTETIQSNANLAPLPPHVPEHLVFDFDMYNPSNLSAGVQEAWAVLQESNVPDLVWTR
                                                      :

P450pyr             YNDIMHVDTNHDIFSSDAGYGGIIIDDG-IQKGGDGGLDLPN---FIAMDRPRHDEQRKA
P450st_1UE8         YEDCKMVLNDHKRFSSNLTGYNDKLEM---LRSGKVFFDIPTRYTMLTSDPPLHDELRNL
CYP119_1F4T         YRYTKEVLNNFSKFSSDLTGYHERLED---LRNGKIRFDIPTRYTMLTSDPPLHDELRSM
P450epok_1PKF       YHDVSAVFRD-ERFAVSREEWESSAEY----SSAIPELSDMKKYGLFGLPPEDHARVRKL
CYP158A2_1S1F       HDDVRLVTND-PRFGR-----EAVMDR---QVTRLAPHFIPARGAVGFLDPPDHTRLRRS
P450nor_1XQD        HKDVCFVATS-EKLSKVRTRQGFPELG------AGGKQAAKAKPTFVDMDPPEHMHQRSM
P450terp_1CPT       HADVMQIGKQPGLFSNAEGSEILYDQNNEAFMRSISGGCPHVIDSLTSMDPPTHTAYRGL
P450cam_1K2O        CNGGHWIATRGQLIREAYEDYRHFSSECP---FIPREAGEAYDFIPTSMDPPEQRQFRAL
                          :      :                                        :   *

P450pyr             VSPIVAPANLAALEGTIRERVSKTLDGLP----VGEEFDWVDRVSIEITTQMLATLFDFP
P450st_1UE8         TADAFNPSNLP--VDFVREVTVKLLSELD------EEFDVIESFAIPLPILVISKMLGIN
CYP119_1F4T         SADIFSPQKLQTLETFIRETTRSLLDSID----P-REDDIVKKLAVPLPIIVISKILGLP
P450epok_1PKF       VNPSFTSRAIDLLRAEIQRTVDQLLDARS----GQEEFDVVRDYAEGIPMRAISALLKVP
CYP158A2_1S1F       VAAAFTARGVERVRERSRGMLDELVDAMLR---AGPPADLTEAVLSPFPIAVICELMGVP
P450nor_1XQD        VEPTFTPEAVKNLQPYIQRTVDDLLEQMKQKGCANGPVDLVKEFALPVPSYIIYTLLGVP
P450terp_1CPT       TLNWFQPASIRKLEENIRRIAQASVQRLLD---FDGECDFMTDCALYYPLHVVMTALGVP
P450cam_1K2O        ANQVVGMPVVDKLENRIQELACSLIESLRP----QGQCNFTEDYAEPFPIRIFMLLAGLP
                        .    :       :      :.            :         .   .     .

P450pyr             FEERRKLTRWSDVTTAAPG-GGVVESWDQRKTELLECAAY----FQVLWNERVNKDPGND
P450st_1UE8         P-DVKKVKDWSDLVALRLGRADEIFSIGRKYLELISFSK-----------KELDSRKGKE
CYP119_1F4T         IEDKEKFKEWSDLVAFRLGKPGEIFELGKKYLELIGYVK-----------DHLNS--GTE
P450epok_1PKF       AECDEKFRRFGSATARALG-VGLVPRVDEETKTLVASVTEGLALLHGVLDERRRNPLEND
CYP158A2_1S1F       ATDRHSMHTWTQLILSSSHGAEVSERAKNEMNAYFSDLIG-----------LRSDSAGED
P450nor_1XQD        FNDLEYLTQQNAIRTNGSSTARQASAANQELLDYLAILVE-----------QRLVEPKDD
P450terp_1CPT       EDDEPLMLKLTQDFFGVEAARRFHETIATFYDYFNGFTVD------------RRSCPKDD
P450cam_1K2O        EEDIPHLKYLTDQMTRPDGSMTFAEAKEALYDYLIPIIEQ------------RRQKPGTD
                          .                                                   :

P450pyr             LISMLAHSPATR-NMTPEEYLGNVLLLIVGGNDTTRNSMTGGVLALHKNPDQFAKLKANP
P450st_1UE8         IVDLTGKIANS--NLSELEKEGYFILLMIAGNETTTNLIGNAIEDFTLYN-SWDYVRE-K
CYP119_1F4T         VVS---RVVNS--NLSDIEKLGYIILLLIAGNETTTNLISNSVIDFTRFN-LWQRIRE-E
P450epok_1PKF       VLTMLLQAEADGSRLSTKELVALVGAIIAAGTDTTIYLIAFAVLNLLRSPEALELVKAEP
CYP158A2_1S1F       VTSLLGAAVGRD-EITLSEAVGLAVLLQIGG-EAVTNNSGQMFHLLLSRPELAERLRSEP
P450nor_1XQD        IISKLCTEQVKPGNIDKSDAVQIAFLLLVAGNATMVNMIALGVATLAQHPDQLAQLKANP
P450terp_1CPT       VMSLLANSKLDGNYIDDKYINAYYVAIATAGHDTTSSSSGGAIIGLSRNPEQLALAKSDP
P450cam_1K2O        AISIVANGQVNGRPITSDEAKRMCGLLLVGGLDTVVNFLSFSMEFLAKSPEHRQELIERP
                              :           :  .*  :         .   :
```

Figure 3.2 (cont. on next page)

```
P450pyr             ALVETMVPEIIRWQTPLAH--MRRTAIADSELGGKTIRKGDKVVMWYYSGNRDDEVIDRP
P450st_1UE8         G-ALKAVEEALRFSPPVMR--TIRVTKEKVKIRDQVIDEGELVRVWIASANRDEEVFKDP
CYP119_1F4T         NLYLKAIEEALRYSPPVMR--TVRKTKERVKLGDQTIEEGEYVRVWIASANRDEEVFHDG
P450epok_1PKF       GLMRNALDEVLRFENILRIG-TVRFARQDLEYCGASIKKGEMVFLLIPSALRDGTVFSRP
CYP158A2_1S1F       EIRPRAIDELLRWIPHRNAVGLSRIALEDVEIKGVRIRAGDAVYVSYLAANRDPEVFPDP
P450nor_1XQD        SLAPQFVEELCRYHTASALA-IKRTAKEDVMIGDKLVRANEGIIASNQSANRDEEVFENP
P450terp_1CPT       ALIPRLVDEAVRWTAPVKS--FMRTALADTEVRGQNIKRGDRIMLSYPSANRDEEVFSNP
P450cam_1K2O        ERIPAACEELLRRFSLVADG---RILTSDYEFHGVQLKKGDQILLPQMLSGLDERENACP
                            *  *           *         .  :  .: :       .  *

P450pyr             EEFIIDRP-RPRQHLSFGFGIHRCVGNRLAEMQLRILWEEILTRFSRIEVMAEPERV---
P450st_1UE8         DSFIPDR--TPNPHLSFGSGIHLCLGAPLARLEARIALEEFAKKFRVKEIVKK-EKI---
CYP119_1F4T         EKFIPDR--NPNPHLSFGSGIHLCLGAPLARLEARIAIEEFSKRFRHIEILDT-EKV---
P450epok_1PKF       DVFDVRR--DTSASLAYGRGPHVCPGVSLARLEAEIAVGTIFRRFPEMKLKETPVFG---
CYP158A2_1S1F       DRIDFER--SPNPHVSFGFGPHYCPGGMLARLESELLVDAVLDRVPGLKLAVAPEDVPFK
P450nor_1XQD        DEFNMNRKWPPQDPLGFGFGDHRCIAEHLAKAELTTVFSTLYQKFPDLKVAVPLGKINYT
P450terp_1CPT       DEFDITR--FPNRHLGFGWGAHMCLGQHLAKLEMKIFFEELLPKLKSVELSGPPRLV---
P450cam_1K2O        MHVDFSR--QKVSHTTFGHGSHLCLGQHLARREIIVTLKEWLTRIPDFSIAPGAQIQH--
                      .   *          :* * * * .  **. :           :.   .:

P450pyr             RSNFVRGYAKMMVRVHA------
P450st_1UE8         DNEVLNGYRKLVVRVERT-----
CYP119_1F4T         PNEVLNGYKRLVVRLKSNE----
P450epok_1PKF       YHPAFRNIESLNVILKPSKAG--
CYP158A2_1S1F       KGALIRGPEALPVTWHA------
P450nor_1XQD        PLNRDVGIVDLPVIF--------
P450terp_1CPT       ATNFVGGPKNVPIRFTKA-----
P450cam_1K2O        KSGIVSGVQALPLVWDPATTKAV
```

Figure 3.2. Partial sequence alignment of P450pyr with members of the P450 family. PDB identification number (% identity with P450pyr, GenInfo Identifier (GI) number): P450st_1UE8 (29%, 51247807), CYP119_1F4T (29%, 11514364), P450epok_1PKF (27%, 38492829), CYP158A2_1S1F (27%, 60593455), P450nor_1XQD (29%, 56554668), P450terp_1CPT (29%, 493949), P450cam_1K2O (24%, 16975144).

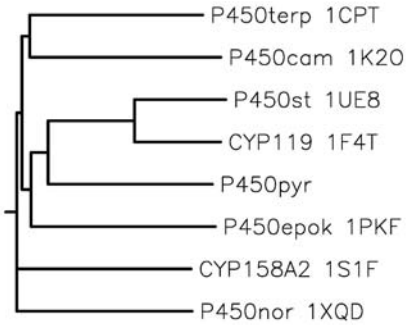


Figure 3.3. Clustal W dendrogram of P450pyr with other members of the P450 family.

Since CYP119 from *Sulfolobus solfactaricus* (1F4T) and P450st from *Sulfolobus tokodaii* (1UE8) shared the highest homology with P450pyr (Figure 3.3), they were used as templates to build a homology model for P450pyr using Insight II software (Insight II, version 2000, Accelrys Inc., San Diego, CA). Using ProStat (Insight II, default parameters), the phi and psi angles were determined to be 82.4%

within their expected values which were comparable to 86.8% and 83.3% for identical analysis of the template PDB structures 1F4T and 1UE8, respectively. A value of 79.9% correct self-compatibility of amino acids with the modeled structure was obtained when inspected by Profiles-3D (Insight II, default parameters). Therefore, the model seemed reasonably well constructed despite the low homology with the templates. *N*-benzyl pyrrolidine **1** was manually docked using the Molecular Operating Environment (MOE, Chemical Computing Group Inc., Montreal, Canada) into the created model and the whole structure was subjected to energy minimization to relieve steric and torsional artifacts from the modeling and docking processes.

Comparison of the P450pyr homology model with X-ray structures of other P450s (Figure 3.4) revealed that all of these P450s have similar conserved "core structures." Two access channels to the catalytic heme site were observed from the surface of the protein (Figure 3.5). The large cavity of the P450pyr active site afforded the possibility for the binding of substrate with bulky protecting groups and contributes to the broad substrate range of this enzyme. The enlarged heme pocket also shows that the ferric ion is located at the corner of the heme pocket. This may contribute the high regioselectivity exhibited in the hydroxylation of *N*-heterocyclic substrates with P450pyr. Within 5Å of the active site are highly conserved residues Thr259, Gly256, and Leu252. The residues Leu251, Leu252, Phe403, and Val404 exhibit a hydrophobic area whereas the polar Gly255 and Gly256 may contribute to inducing the orientation of the hydrophilic group in the substrate.

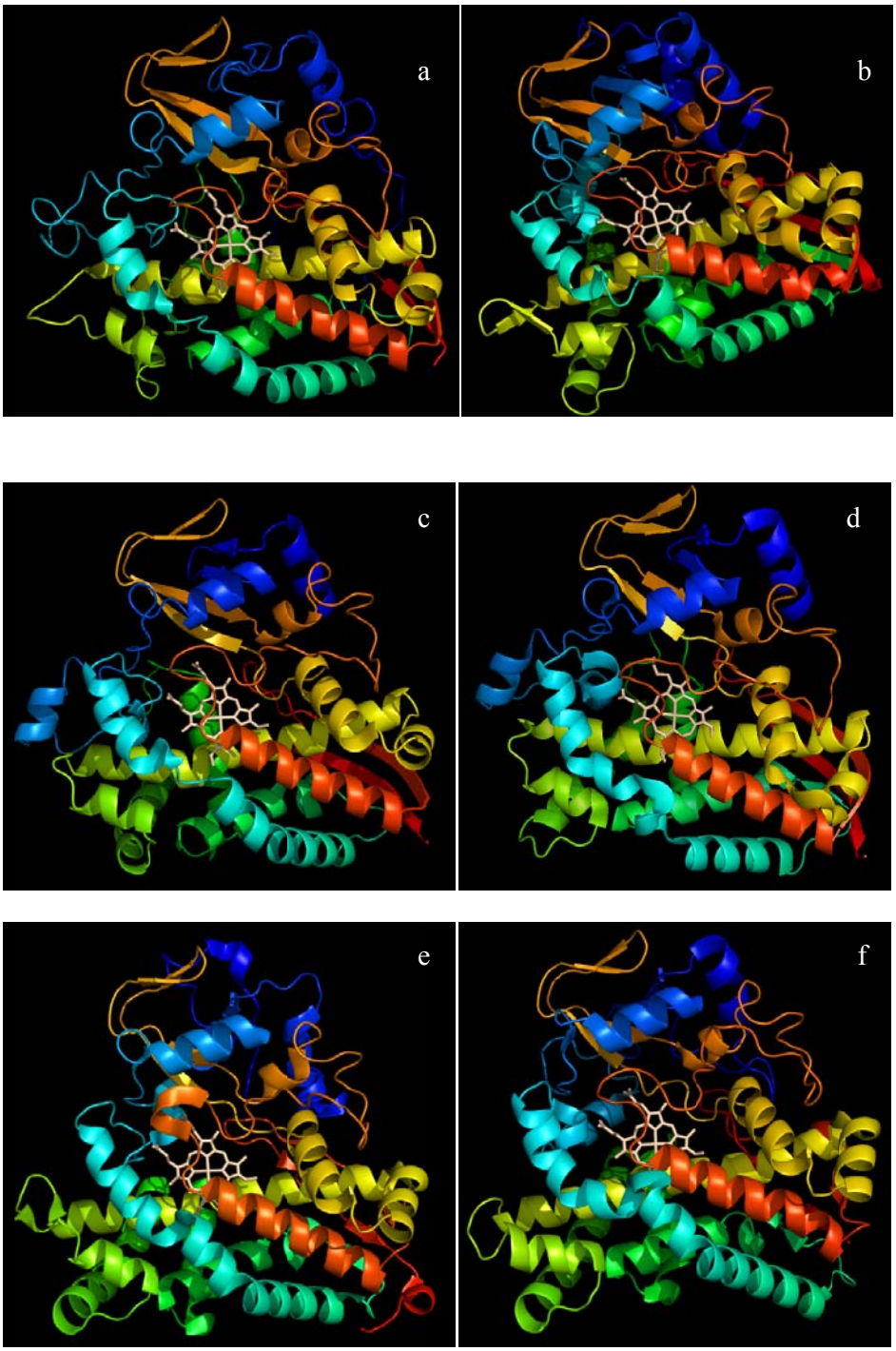


Figure 3.4. Structure comparison of P450pyr (a) with P450terp (b), CYP119 (c), P450st (d), P450cam (e), and P450nor (f).

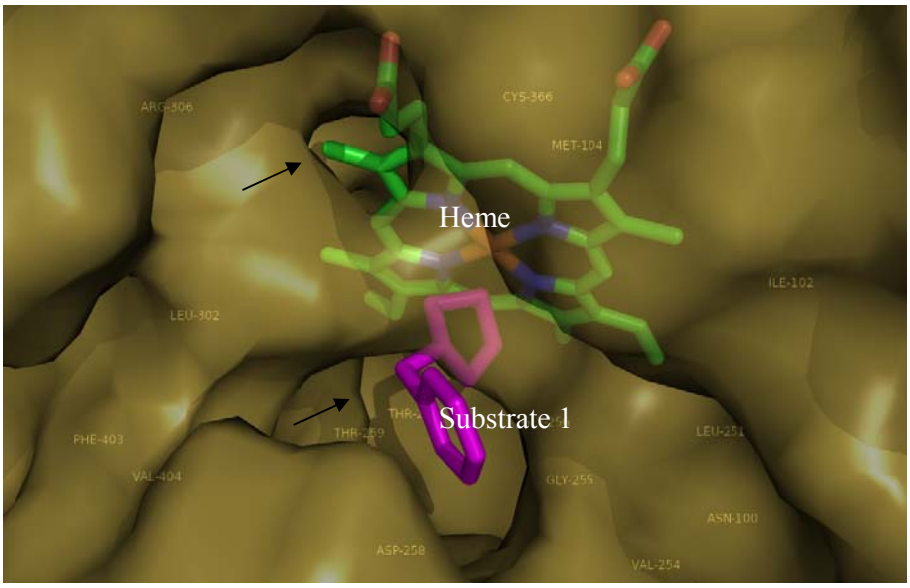


Figure 3.5. Surface around the P450pyr active site. Substrate **1** is shown in magenta. The ferric ion on the heme is displayed in red. The black arrows point to the two access channels to the catalytic heme site of P450pyr.

3.2.2 Cloning and Expression of Cytochrome P450pyr Electron Transport System

Whole-cell assay was necessary in this case as our P450pyr system requires the close interaction of the ferredoxin and ferredoxin reductase to form an electron transport system, unlike P450BM3 which has its hydroxylase and reductase domains on a single polypeptide. Also, the whole-cell assay approach removed the need for the cell lysis step, and the addition of NADH and other auxiliary protein components, thus simplifying our screening protocol. Previously, the *Pseudomonas putida* strain expressing the P450pyr together with the ferredoxin (Fdx) gene from *Sphingomonas* sp. HXN200 and the ferredoxin reductase (FdR) gene from *Mycobacterium* sp. HXN1500 was developed but showed low activity.[131] *E. coli* was chosen as the host strain due to its high growth rate, the abundant availability of genomic tools, and the

relative completeness and availability of its genetic information. Initially, we cloned the genes encoding the P450pyr, Fdx, and FdR into different Duet expression vectors (i.e. pACYCDuet, pETDuet and pRSFDuet) and utilized a whole-cell assay approach to determine the substrate conversion percentage of each construct (Table 3.1). The recombinant *E. coli* BL21(DE3) strain containing a dual plasmid system (pRSFDuet P450pyr and pETDuet Fdx FdR; Figure 3.6) showed the highest hydroxylation activity and was thus designated as the parent strain. The balance between the expression level of P450pyr and the Fdx and FdR genes plays an important role in determining the activity. A clear band at 48 kDa indicating the P450pyr enzyme was observed on an SDS-PAGE gel of the cell extract of the parent strain (Figure 3.7).

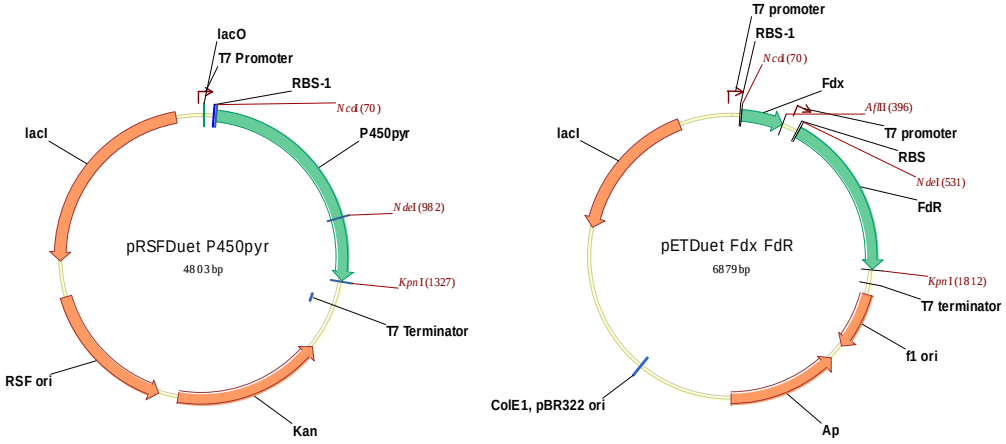


Figure 3.6. pRSFDuet P450pyr and pETDuet Fdx FdR expression vector. The P450pyr gene was cloned into the *NcoI*/*KpnI* site of pRSFDuet (left vector) whereas the ferredoxin Fdx gene and the ferredoxin reductase FdR gene was cloned into the *NcoI*/*AflII* site (MCS1) and *NdeI*/*KpnI* site (MCS2) of pETDuet (right vector), respectively. These two plasmids were then transformed into BL21(DE3).

Table 3.1. Conversion of substrates *N*-Benzyl-pyrrolidine 1 and *N*-benzyloxycarbonyl-pyrrolidine using a whole-cell system.

Whole-cell system	***N*-Benzylpyrrolidine Conversion (30 min)**	***N*-benzyloxycarbonyl-pyrrolidine Conversion (30 min)**
BL21 pRSFDuet P450pyr pETDuet FR	39.2%	88.3%
BL21 pACYCDuet P450pyr pETDuet FR	23.8%	87.8%
BL21 pACYCDuet P450pyr pRSFDuet FR	(cells took more than 18 h to grow to saturation)	
BL21 pETDuet P450pyr pACYCDuet FR	ND	1%
BL21 pETDuet P450pyr pRSFDuet FR	ND	8%
BL21 pRSFDuet P450pyr pACYCDuet FR	ND	1%
JM109 pCom8 P450pyr FR	0.5%	6.3%
BL21 pCom8 P450pyr FR	1.9%	ND

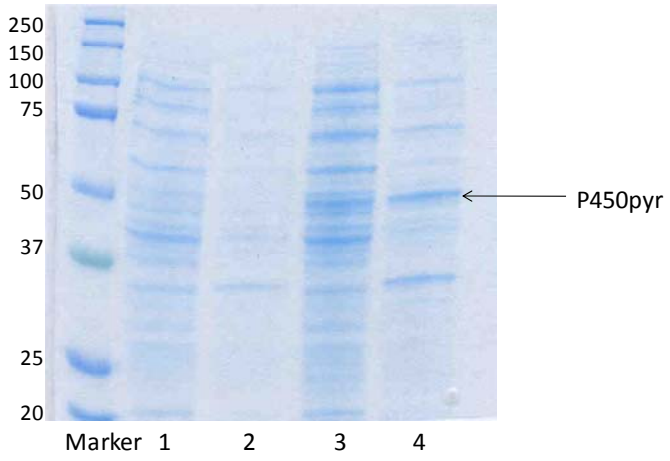


Figure 3.7. SDS-PAGE analysis of control BL21(DE3) pRSFDuet pETDuet (empty vectors) soluble (lane 1) and insoluble (lane 2) fractions, as well as BL21 pRSFDuet P450pyr pETDuet FR soluble (lane 3) and insoluble fractions (lane 4). The P450pyr enzyme was observed as a clear band at 48 kDa.

3.2.3 Iterative Targeted Site Saturation Mutagenesis

The mutant library was created using iterative targeted site-saturation mutagenesis.[37,40,173] The strategy can be described as follows: i) identify all residues which are in the vicinity of the heme-docked substrate, ii) perform individual site saturation mutagenesis of all or a subset of these selected residues, iii) screen each library in 96-well plates, iv) select mutants with improved

enantioselectivity over wild type (WT), v) perform another round of individual site saturation mutagenesis at the remaining unmutated selected residues, and vi) repeat steps iii-v until no further improvement can be achieved.

Based on the P450pyr homology model, 17 residues were identified within 5Å of the heme-docked substrate (Fig 3.8a) and were subjected to individual site saturation mutagenesis. C366 and G256 were not subjected to randomization because they are highly conserved residues involved in the heme coordination and catalytic activity of the P450pyr, respectively. For each site, 32 distinct variant possibilities exist (32 possible codon substitutions). Hence, the screening of 180 transformants per site should be more than enough to provide comprehensive coverage of all created variants. This protein engineering approach involves screening very manageable size of mutant libraries which target the residues within the substrate binding pocket.

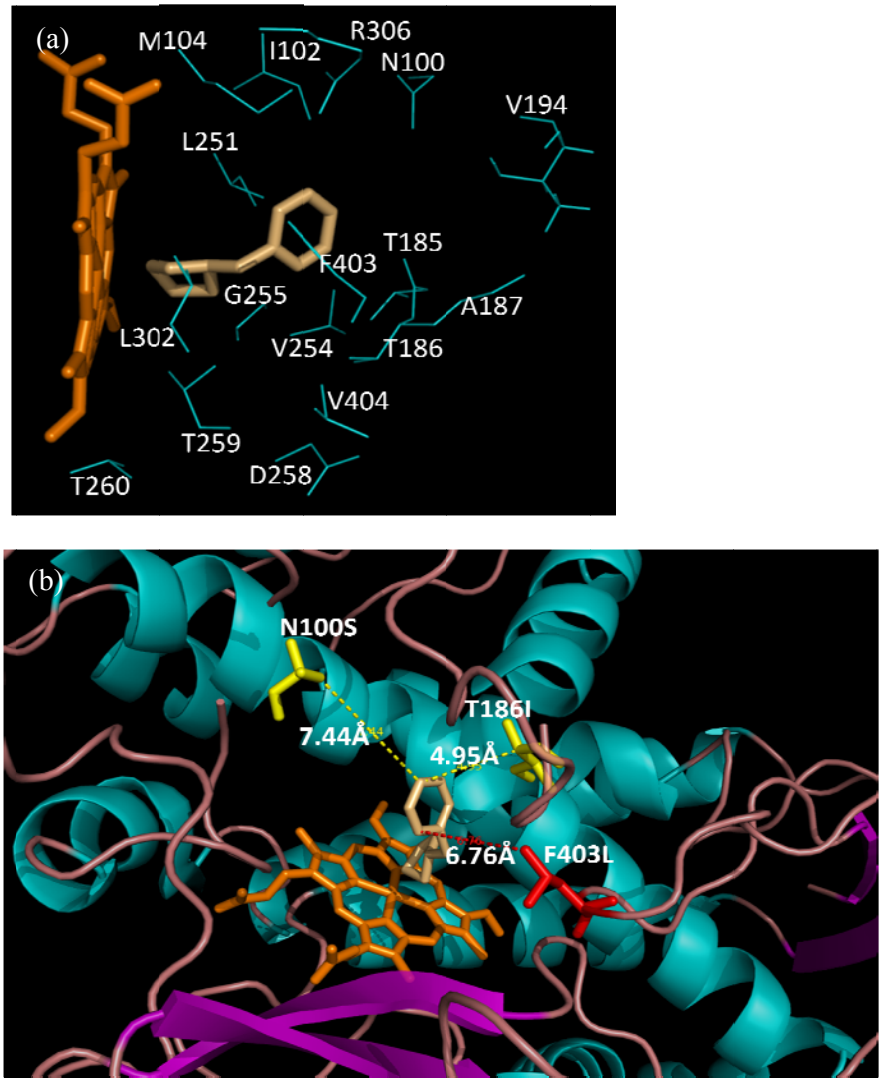


Figure 3.8. (a) Homology model showing the 17 residues that were identified within 5Å of the heme (orange)-docked substrate (light orange). (b) The mutation sites are shown and labeled (yellow for the (*R*)-mutant and red for the (*S*)-mutant), and the distances between the substrate 1 and the mutation site are shown in Å.

3.2.4 Screening Strategy

The development of an efficient high-throughput enantiomeric excess (*ee*)-screening method was no trivial task. The high-throughput two-enzyme-based colorimetric assay (Chapter 2) was applied to semi-quantitatively determine the enantiomeric ratio of the racemic product.

Our high-throughput screening protocol can be described as following: Firstly, the mutant libraries were inoculated and expressed in TB medium in 96-deep well plates, in the presence of δ-aminolevulinic acid (δ-ALA) to increase the expression level of P450pyr. The biohydroxylation assay to convert **1** to (*R*)- and (*S*)-**2** was performed with resting cells in 96-deep well plates with shaking at 1100 rpm to facilitate better mixing and mass transfer. When the biohydroxylation reaction was carried out at low shaking speed of 250 to 500 rpm, negligible substrate conversion was observed. After the biohydroxylation with whole-cells, the deep well plates were centrifuged and aliquots (80 μL) of the supernatant containing the unreacted **1** and its racemic product **2** were pipetted into two separate 96-well microtiter plates. In each plate, purified BRD or RDR and its corresponding cofactor were added. By utilizing the nitro blue tetrazolium (NBT) - phenazine methosulfate (PMS) colorimetric assay[122], the production of NAD(P)H was monitored by observing the formation of partially soluble purple formazan at OD_{580} which corresponded to the activity of the dehydrogenases that in turn is correlated to the concentration of each enantiomer in the racemic product (Figure 3.9). Hence, the estimated *ee* of each mutant could be compared against the WT P450pyr. Mutants that showed a greater *ee* than that of the WT were selected and subjected to a larger-scale biohydroxylation assay in shaking flasks to verify the enantioselectivity of the mutants by chiral HPLC analysis.

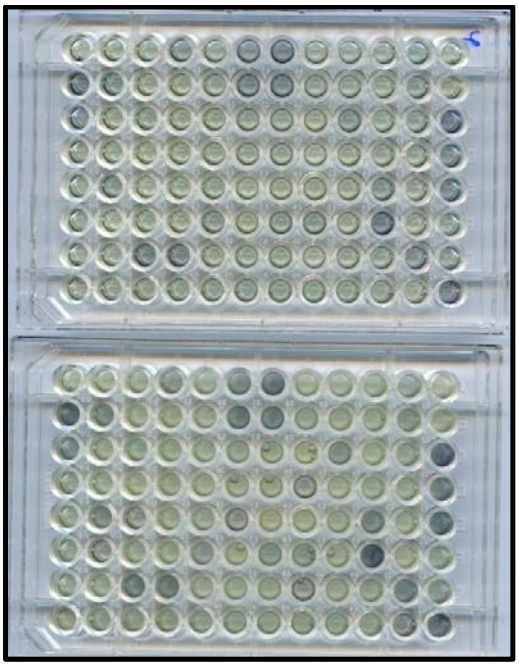


Figure 3.9. Example of 96-well microtiter plate screening using the BRD and RDR enzymes coupled with the NBT-PMS assay. The formation of purple formazan which correlates to the concentration of (*S*)- and (*R*)-**2** is apparent even to the naked eye and its intensity can be measured using the microplate reader.

In the first round of saturation mutagenesis, using the WT P450pyr as a template, we screened 180 mutants per site, which translates into a total of 3060 mutants (17 sites) screened. Two interesting mutants with preference for (*S*)- and (*R*)-**2** were identified (Table 3.2). The (*S*)-selective mutant 11BB12 (F403L) displayed an *ee* of 65% which was a 1.5 times improvement over that of the WT P450pyr. Mutant 1AF4 (N100S), on the other hand, displayed a complete reversal of enantioselectivity with an *ee* of 42% with a preference for (*R*)-**2**. In fact, it was identified twice in the screening of the same 3060-mutant library. It was intriguing that only one mutational change at position 100 (replacing asparagine with serine) could induce a total inversion of the P450pyr enantioselectivity.

Encouraged by the success in the first round, F403L was used as a template for a second round of saturation mutagenesis on the remaining 16 residue sites.

Unfortunately, we failed to discover any (*S*)-selective mutants with significantly higher *ee* than the parent 11BB12. Hence, we decided to identify more (*R*)-selective mutants instead. Using N100S as a template, a third round of saturation mutagenesis was conducted on the remaining 16 residue sites. In this round, we discovered four mutants with improved (*R*)-enantioselectivity ranging from *ee* of 60% to 83%. The best mutant, 1AF4A, was sequenced and it was found that the mutational change at position 186 (isoleucine substitution of threonine) was responsible for the further increase in the (*R*)-enantioselectivity of the mutant P450pyr.

Table 3.2. Hydroxylation of *N*-benzyl pyrrolidine **1** by engineered cytochrome P450pyr variants.

Description	**Mutant**	**Amino Acid Changes**	**Conversion (%)**[a]	***ee* (%)**[b]
WT P450pyr	NA	NA	55	43 (*S*)
Round 1	1AF4	N100S	33	42 (*R*)
	11BB12	F403L	47	65 (*S*)
Round 2	No significantly improved mutants			
Round 3	1AF4A	N100S, T186I	23	83 (*R*)
	1AF4B	N100S, T259S	16	63 (*R*)
	1AF4C	N100S, L302V	19	71 (*R*)
	1AF4D	N100S, V404K	27	60 (*R*)

[a] Conversion determined based on substrate consumption with whole-cell assay (10 g cell dry weight (cdw) per liter) using starting substrate concentration of 5 mM. Biohydroxylation time was 4 h.
[b] Determined by chiral HPLC analysis using the Chiralcel OB-H (250 mm x 4.6mm) column. All verification experiments were carried out in shaking flasks in triplicates.

The conversion of **1** by the P450pyr mutants is comparatively lower than that of the WT P450pyr (Table 3.2). The conversion of the best mutant 1AF4A was more than 50% lower than the WT P450pyr. The introduction of the mutation in the active pocket of the P450pyr has probably introduced a destabilizing effect in the enzyme and thus negatively affecting its activity, although the enantioselectivity of the enzyme has been significantly improved. For future work, we could perhaps take a step back and apply error-prone PCR (epPCR) to the full-length gene encoding the 1AF4A mutant to improve its activity to a level equal to

or higher than the WT P450pyr.

3.2.5 Combination of Beneficial Mutations by Site Directed Mutagenesis

To evaluate the effects of combined mutations, variants with triple mutations by combining all the mutations found in Round 3 (Table 3.2) were constructed by site directed mutagenesis. Unfortunately, the introduction of these mutations rendered the P450pyr enzyme reduced its activity significantly (Table 3.3). Although the *ee* was still reasonably high, ranging from 66 to 81%, there was no improvement over our best mutant 1AF4A. Also, the low hydroxylation activity renders the mutants not feasible for practical applications.

Table 3.3. Effect of combination of mutations on substrate conversion and *ee*

Mutation sites	Conversion (%)	*ee* (%)
N100S T186I V404K	7	66 (*R*)
N100S T186I T259S	6	81 (*R*)
N100S T259S L302V	6	81 (*R*)
N100S T259S V404K	5	78 (*R*)
N100S L302V V404K	<1	ND
N100S T186I L302V	<1	ND

[a] Conversion determined based on substrate consumption with whole-cell assay (10 g cdw per liter) using starting substrate concentration of 5 mM. Biohydroxylation time was 4 h.
[b] Determined by chiral HPLC analysis using the Chiralcel OB-H (250 mm x 4.6mm) column. ND: Not determined

3.3 Discussion

3.3.1 Evolutionary Strategy

In this chapter, we described the application of a general protein engineering approach that we have applied to reengineer the enantioselectivity of the P450pyr monooxygenase. We hypothesized that changes in enantioselectivity of the enzyme are heavily influenced by the interactions between the substrate and the residues in the

active binding pocket. By combining stepwise targeted site saturation mutagenesis of the active site protein residues which are within 5Å of the heme-docked substrate, we have been able to modify the P450pyr enantioselectivity for substrate **1.** Mutants with improved (*R*)- and (*S*)- selectivity over the WT has been verified and identified.

The stepwise individual site saturation mutagenesis strategy has the following advantages: (i) Universality: The current library creation strategy can be readily generalized to engineer the enantioselectivity of other proteins. Only the homology model of the protein is required. (ii) Screenability: There are only 19 possible amino acid substitutions or 32 possible codon substitutions per site for the saturation mutagenesis libraries. Thus, screening two 96-well plates should be sufficient to represent all possible library variants. The library size makes it easier to screen and less labor intensive. (iii) Efficiency: Due to the small library size of protein mutants, all randomized mutants can be subjected to simultaneous screening for the (*R*)- and (*S*)-mutants in a faster yet comprehensive manner. These focused libraries are also of higher quality, provided the sites to be randomized have been chosen correctly. It minimizes the undesired formation of inferior or inactive enzyme mutants and thus reduces the overall screening effort, which can be substantial in the case of enantioselectivity.[132] (iv) Accessibility to amino acid substitutions: Stepwise site saturation mutagenesis approach allows every site in the active site pocket to be randomized to all 20 possible amino acids, ensuring that important amino acid substitutions at critical binding positions are not overlooked.

We should also note that the mutants which were identified using the stepwise saturation mutagenesis need not represent the only variants with enhanced enantioselectivity. The initial template for each round of saturation mutagenesis

would influence the next round of mutagenesis and screening. If another mutant were chosen as the starting point for the subsequent round of screening, the resultant final variant might be a different one.

3.3.2 Structural Analysis of Mutations

The roles played by individual mutations towards altering the enantioselectivity of the P450pyr were not clear. The inversion of enantioselectivity can arise by the insertion of the *O*-atom either from the opposite direction into the C-H bond at the original 3-position or from the same direction into the C-H bond at the adjacent position due to the special C(2v) symmetry nature of the pyrrolidine substrate. From the homology model, a large cavity allowing substrate access to the heme site was observed, with hydrophobic residues lining the entrance to the cavity. An interesting observation was that the mutation sites 100, 403, and 186 were located near the entrance of the cavity, suggesting that a benzyl ring-protein interaction furthest from the heme may induce enantioselectivity towards **1** (See Figure 3.8b). The single amino acid substitution of the Asn100 residue by Ser that caused a complete inversion of the P450pyr enantioselectivity is of great interest. We reasoned that changing the bigger hydrophilic Asn residue to the smaller neutral Ser residue would increase the active-site volume and alter the conformation of the helix containing the N100S upon substrate binding, which may be responsible for the reversal of enantioselectivity. When the neutral Thr residue was also mutated to a hydrophobic Ile, the cooperative effect of the two mutations increased the enantioselectivity towards the (*R*)-enantiomer by almost 2 fold. Interestingly, in both cases (N100S and T186I), the residue was replaced by a more hydrophobic residue. The Phe403 in the WT P450pyr is located in the middle of the cavity to the active site

pocket, thus providing a hydrophobic barrier for the hydrophilic directing groups in incoming substrates. In the case of the (*S*)-enantioselective mutant, 11BB12, the aromatic hydrophobic side chain of Phe was replaced by the aliphatic side chain of Leu, thus slightly reducing the hydrophobicity which may have directly or indirectly increased the mutant's enantioselectivity towards the (*S*)-enantiomer due to conformational changes at entrance to the enzyme's active site.

3.4 Conclusions and Outlook

In summary, we have inverted the enantioselectivity of the P450pyr towards the biohydroxylation of prochiral substrate **1** by iterative targeted site-saturation mutagenesis coupled with a high-throughput two-enzyme-based colorimetric assay. Interestingly, one mutation at position 100 was all it took to completely invert the enantioselectivity of the P450pyr. The best mutant 1AF4A has an *ee* of 83% (*R*) compared to the wild type's *ee* of 43% (*S*). This inversion was achieved in only two rounds of saturation mutagenesis with two mutation sites. In addition, we have also discovered the (*S*)-enantioselective mutant 11BB12 that has an *ee* of 65% (*S*). This protein engineering approach involves screening very manageable size of mutant libraries which targets the residues within the substrate binding pocket. We have also demonstrated the first example in improving the biohydroxylation enantioselectivity of a class I P450 monooxygenase via a directed evolution approach combined with an *ee* high-throughput screening assay.

Perhaps it would also be beneficial if we took a step back to first improve the P450pyr stability before proceeding with further rounds of mutagenesis to improve the *ee*. Extra stability could perhaps confer more evolvability as the stabilized P450 parent could tolerate the highly destabilizing mutations.[174]

3.5 Materials and Methods

Materials

Escherichia coli BL21(DE3) and the expression plasmids were purchased from Novagen (Madison, WI). PCR grade dNTPs were obtained from Roche Applied Sciences (Indianapolis, IN). Restriction enzymes, Phusion High-Fidelity DNA polymerase, T4 DNA ligase, and their corresponding buffers were purchased from New England Biolabs (NEB) (Beverly, MA). D-Glucose was purchased from ThermoFisher (Pittsburgh, PA). Ampicillin, kanamycin, isopropyl *β*-D-thiogalactopyranoside (IPTG), nitro blue tetrazolium (NBT), phenazine methosulfate (PMS), NAD^+, and $NADP^+$ were purchased from Sigma (St. Louis, MO). Other required salts and reagents were purchased from either Fisher or Sigma-Aldrich. The Ni-NTA agarose, QIAprep spin plasmid mini-prep kit, QIAEX II gel purification kit, and QIAquick PCR purification kit were purchased from Qiagen (Valencia, CA). Various oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA). SDS-PAGE gels, buffers and protein size markers were purchased from Bio-Rad (Hercules, CA).

Homology Modeling

The following structures were downloaded from the Protein Data Bank (PDB) database (PDB accession code): CYP119 from *Sulfolobus solfactaricus* (1F4T) [175] and P450st from *Sulfolobus tokodaii* (1UE8)[176]. These two structures were aligned using Insight II software (Insight II, version 2000; Accelrys Inc., San Diego, CA) based on its conserved structural regions to achieve the lowest root-mean-square (RMS) score.

The amino acid sequence of P450pyr was manually aligned with the structural alignment such that the aligned sequences represented homologous structural regions. Using default parameters of the automated MODELER module within Insight II, with the alignment as input and using moderate refinement of the structure and loop regions, we obtained approximately 30 structural models. By visually inspecting the models for obvious flaws, comparing the scores from the Profiles 3-D function, and the ProStat inspection of ϕ and ψ angles, the best model was selected. Docking of substrate **1** onto the heme binding site of the best model was carried out using the Molecular Operating Environment (MOE; Chemical Computing Group Inc., Montreal, Canada). The whole structure was then subjected to energy minimization to relieve steric and torsional artifacts from the modeling and docking processes.

Engineering of a Recombinant *E. coli* Strain Expressing the Active P450pyr

We were interested in engineering a recombinant *E. coli* strain expressing the P450pyr and ferredoxin (Fdx) genes from *Sphingomonas* sp. HXN200 and the ferredoxin reductase (FdR) gene from *Mycobacterium* sp. HXN1500. Plasmid pCom8-PA7F200R1500 [131] was used as a PCR template to amplify the coding region of cytochrome P450pyr, ferredoxin and ferredoxin reductase. The P450pyr gene was amplified using the forward primer P450fwd (5' TTA ACT ACT CCA TGG AAC ATA CAG GAC AAA GCG CGG 3', the *Nco*I restriction site is underlined) and reverse primer P450rev (5' TTA ACT ACT GGT ACC CTA CGC GTG GAC GCG AAC 3', the *Kpn*I restriction site is underlined). The Fdx gene was amplified using the forward primer Frdxfwd (5' TTA ACT ACT CCA TGg ca ACA GTG ACC TAT GTT GAA ATA AAT GGC AC 3', the *Nco*I restriction site is underlined, the second codon of the Fdx gene, in small caps, was changed from the original cca to gca) and

reverse primer Frdxrev (5' TTA ACT ACT CTT AAG TCA ATG CTG CGC GAG AGG AAG 3', the *Afl*II restriction site is underlined). The FdR gene was amplified using the forward primer R1500fwd (5' TTA ACT ACT CAT ATG ATC CAC ACC GGC GTG AC 3', the *Nde*I restriction site is underlined) and the reverse primer R1500rev (5' TTA ACT ACT GGT ACC TTA GAG GGA GGT TGG GGA CG 3', the *Kpn*I restriction site is underlined). PCR amplified genes were digested with their respective enzymes. The digested P450pyr fragment was ligated into the pRSFDuet vector that was digested with *Nco*I and *Kpn*I restriction enzymes. The digested Fdx fragment was ligated into the multiple cloning site (MCS) 1 of the pETDuet vector that was digested with *Nco*I and *Afl*II. The resulting plasmids from the ligations were separately transformed into chemical competent *E. coli* DH5α cells which were then plated on Luria-Bertani (LB) agar plates containing 50 μg/mL kanamycin (for pRSFDuet P450pyr) and 100 μg/mL ampicillin (for pETDuet Fdx), respectively. Subsequently, the digested FdR fragment was ligated into the MCS 2 of the same pETDuet vector, containing the Fdx gene, digested with *Nde*I and *Kpn*I restriction enzymes. The resulting plasmid from the ligation was transformed into chemical competent *E. coli* DH5α cells which were then plated on LB agar plates containing 100 μg/mL ampicillin. Finally, the purified pRSFDuet P450pyr and pETDuet Fdx plasmids were transformed into electrocompetent *E. coli* BL21(DE3) and plated on LB agar plates containing 50 μg/mL kanamycin and 100 μg/mL ampicillin.

Library Generation

Megaprimer PCR method was used to generate libraries where single residues were randomized to all 20 possible amino acids. In the first round of saturation

mutagenesis, pRSFDuet P450pyr wild type plasmid was used as template. Briefly, three primers were used to generate an amplified gene library consisting of a saturation mutagenized residue: two primers flanking the P450pyr region (PF200fwd, 5'-TTA ACT ACT CCA TGG AAC ATA CAG GAC AAA GCG CGG-3', and P450rev KpnI, 5'-TTA ACT ACT GGT ACC CTA CGC GTG GAC GCG AAC-3') and a degenerate primer incorporating the residue to be mutated. The degenerate primer substituted the codon corresponding to the target residue with the sequence NNS and contained 9-10 additional bases on either side (5' and 3'). The NNS substitution allowed the incorporation of all 20 amino acids while limiting the codon possibilities to only 32. For each gene library with a randomized codon, 2 PCRs were performed. Firstly, the megaprimer containing the NNS substitution at the codon of interest was amplified. Each PCR contained (50 mL of final volume): 1X Phusion HF buffer, 0.2 mM of each dNTP (Roche), 0.5 μM of appropriate degenerate primer, 0.5 μM of appropriate flanking primer (PF200fwd or P450rev KpnI), and 50 ng of template plasmid. PCRs were carried out on a MJ Research (Watertown, MA) PTC-200 thermocycler with cycle conditions of 98 °C 30 s, (98 °C 10 s, 58 °C 30 s, and 72 °C 15 s) × 25 cycles and 72 °C 10 min. The megaprimer product was purified and eluted into 50 μL EB buffer using the QIAquick PCR purification kit. 10 μL of the megaprimer was then combined into a 100 μL PCR reaction which shares the same conditions to the first PCR step described above, except that the primers used were replaced with the megaprimer and its corresponding flanking primer, and a different program was used: 98 °C 30 s, (98 °C 10 s, 58 °C 30 s, and 72 °C 45 s) × 30 cycles and 72 °C 10 min. PCR product from this reaction was purified from a 1% agarose gel by using the QIAEX II gel purification kit. The insert was digested with *Nco*I and

*Kpn*I and ligated into the pRSFDuet vector. This library was electroporated into *E. coli* BL21(DE3) competent cells containing the pETDuet Fdx FdR vector.

Library Screening

For identification of enantioselective P450pyr variants, a 96-well plate two-enzyme based colorimetric assay was developed based on a colorimetric nitro blue tetrazolium (NBT) - phenazine methosulfate (PMS) [122]. Library colonies were picked with sterile toothpicks and used to inoculate 500 µL of TB medium containing 50 µg/mL kanamycin and 100 µg/mL ampicillin in 96-deep well (2 mL) plates. The plates were shaken at 500 rpm, 37 °C for 6 h, and protein expression was induced by adding 500 µL of TB medium containing 0.5 mM IPTG and 0.5 mM δ-ALA, followed by shaking at 500 rpm, 30 °C for 14 h. The plates were centrifuged at 4000 rpm for 10 min at 4 °C, and the cells were resuspended in 300 µL of 50 mM potassium phosphate buffer pH 7.0 containing 2 % (w/v) D-glucose and 10 mM of **1**. The plates were then shaken at 1100 rpm, 30 °C for 5 h and were later centrifuged at 4000 rpm for 10 min at 4 °C. An aliquot (80 µL) of the supernatant was transferred to two fresh plates, one containing ~5 µg BRD enzyme while the other plate contained ~5 µg of the RDR enzyme. The NBT assay solution (final concentration 2 mg/mL NBT, 0.1 mg/mL PMS, 0.5 mM NAD(P)H) was added to each plate to initiate the reaction. The plates were incubated at room temperature in darkness for 2.5 h and the absorbance at 580 nm was measured in a SpectraMax 340PC microplate reader (Molecular Devices, Sunnyvale, CA). Enantioselective mutants were identified by comparing the ratio of the absorbance on both plates against that of the parent enzyme.

Whole-cell Assay in Shaking Flasks

To confirm the positive mutants, they were streaked on LB plates containing 50 μg/mL kanamycin and 100 μg/mL ampicillin, and single colonies were picked into 5 mL of liquid LB medium containing the antibiotics. After shaking at 37 °C overnight, 1 mL of the culture were inoculated into 10 mL of liquid TB medium containing the antibiotics, 0.5 mM IPTG and 0.5 mM δ-ALA. The culture was shaken at 250 rpm, 30 °C for 6 h and then centrifuged at 4000 rpm for 5 min at 4 °C. The cells were resuspended (to 10 g cell dry weight per liter) in 50 mM potassium phosphate buffer pH 7.0 containing 2 % (w/v) D-glucose and 10 mM of **1**, and further shaken at 250 rpm, 30 °C for 4 h. After centrifuging at 4000 rpm, 80 μL aliquot of the supernatant was taken to be analyzed by using the Hypersil BDS-C18 (5 μm) column (15 mm × 4 mm) on the reverse-phase HPLC [130]. The chiral product in the remaining supernatant was extracted using ethyl acetate in a standard work up procedure and the chirality was analyzed by using the Chiralcel OB-H (250 mm x 4.6mm) column [126]. The *ee*'s of the positive mutants were verified in triplicates.

Site Directed Mutagenesis

Overlap extension PCR method was used introduce mutations at desired locations. As an example, to introduce 3 different mutations, the mutant already containing 1 of the mutation was used as a template so that we only need to introduce another 2 more mutations. As illustrated in Figure 3.10, six primers were used: two primers (aF and dR) flanking the template and 2 pairs of complementary degenerate primers (bF and bR, cF and cR) incorporating the residue to be mutated. The degenerate primer substituted the codon corresponding to the target residue with the mutation desired and contained 9-10 additional bases on either side (5' and 3'). For

each triple mutant, only 2 PCR rounds need to be performed. In PCR Round 1, three fragments: AB, BC and CD, containing the desired mutation were amplified using the forward and reverse primer pairs aF and bR, bF and cR, as well as cF and dR, respectively. Each PCR contained (50 mL of final volume): 1× Phusion HF buffer, 0.2 mM of each dNTP (Roche), 0.5 μM of the primer pairs, and 50 ng of template plasmid. PCRs were carried out on a MJ Research (Watertown, MA) PTC-200 thermocycler with cycle conditions of 98 °C 30 s, (98 °C 10 s, 58 °C 30 s, and 72 °C 15 s) × 25 cycles and 72°C 10 min. The PCR amplified fragments from this reaction was purified from a 1% agarose gel and eluted into 30 μL EB buffer using the QIAEX II gel purification kit. In PCR Round 2, ~50ng of each fragment and the template flanking primers, aF and dR, were then combined into a 100 μL PCR reaction which shares the same conditions to the first PCR step described above, except the template was replaced with the 3 fragments and a different program was used: 98 °C 30 s, (98 °C 10 s, 58 °C 30 s, and 72 °C 45 s) × 30 cycles and 72 °C 10 min. The PCR product was PCR purified using the QIAquick PCR purification kit, digested with *Nco*I and *Kpn*I and ligated into the pRSFDuet vector. This triple mutant was then electroporated into *E. coli* BL21(DE3) competent cells containing the pETDuet Fdx FdR vector and subsequently subjected to a whole-cell assay to determine its substrate conversion % and *ee*.

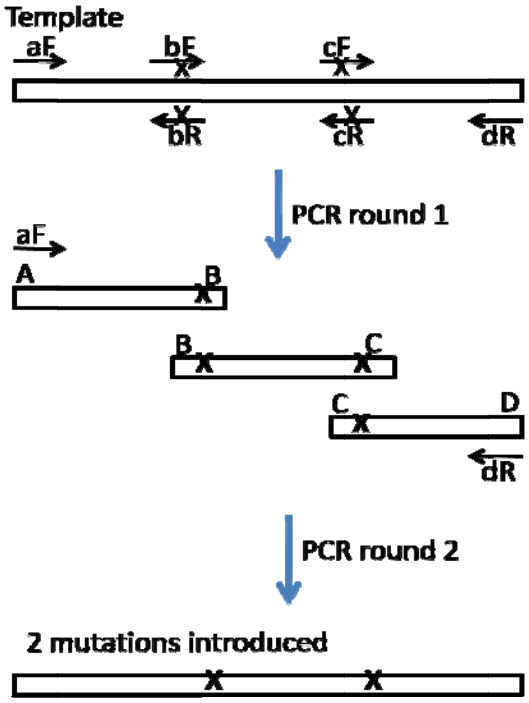


Figure 3.10. General scheme of the overlap extension PCR method that was used to introduce site-directed mutations to the template.

DNA Sequencing and Analysis

Plasmid DNA from selected mutants was isolated using QIAprep spin plasmid mini-prep kits. Sequencing reactions consisted of 100 to 200 ng of template DNA, 10 pmol of each primer, sequencing buffer, and the BigDye reagent (Applied Biosystems, Foster City, CA). Reactions were carried out for 31cycles of 96 °C for 20 s, 50 °C for 10 s, and 60 °C for 3.5 min in an MJ Research PTC-200 thermal cycler (Watertown, MA). Prepared samples were submitted to the Biotechnology Center at the University of Illinois for sequencing on an ABI PRISM 3700 sequencer (Applied Biosystems, Foster City, CA).

Chapter 4 : Development of a Simple, Efficient and General Method for Cofactor Recycling in a Bio-Oxidation

4.1 Introduction

The majority of the enzymes that are now being developed or are currently utilized in industry are mostly hydrolytic in nature.[65-67] These enzymes may be applied in the form of whole-cells or as isolated enzymes in a soluble or immobilized state. Classes of enzymes such as hydrolases, isomerases or lyases are cofactor-independent, or have a cofactor like a metal ion or a coenzyme that is firmLy bound. In contrast, enzymes such as oxidoreductases, transferases and ligases are cofactor-dependent and have been found to perform more complex chemistry that are synthetically useful.[67-72] For example, NAD(P)H-dependent cytochrome P450s catalyze hydroxylation, *N*-, *O*-, *S*-dealkylation, sulphoxidation, epoxidation, deamination, desulphuration, dehalogenation, peroxidation and *N*-oxide reduction[12], NAD-dependent oxidoreductases catalyze the asymmetric reduction of carbonyl groups to alcohols and amines[72,73] and coenzyme A-dependent (CoA) synthetases catalyze asymmetric carbon-carbon bond formation.[74-77] However, as these enzymes use expensive cofactors, they have not been as widely applied in industrial settings compared to cofactor-independent enzymes.

In a cell, the pyridine nucleotide cofactors NAD^+/NADH and $NADP^+$/NADPH are essential components as they function as electron carriers in reduction and oxidation reactions. Whole-cell biocatalysis provide a continuous source of enzymes and cofactors, and is thus generally preferred over purified enzymes in industrial settings, especially for cofactor-dependent biotransformations. Although it is known

that whole-cells have some reserves of cofactors, cofactor depletion is often a limiting factor in whole-cell-catalyzed reactions. Thus, in order for a cofactor-dependent reaction to be practical, an efficient recycling of the cofactor is necessary.

Strategies for nicotinamide cofactor regeneration can be classified into four general categories: biological, enzymatical, electrochemical, chemical and photochemical.[79-87] The enzymatical strategy can be further divided into the coupled-substrate[81,177-179] and the coupled-enzyme approaches.[123,180-187] In the former, only one enzyme is used to catalyze both the main and the regeneration reactions. This method is suitable for systems that use enzymes in organic media because the organic solvent is used as a substrate. The drawback of this method is that the choice of the co-substrate is limited by the enzyme's substrate specificity. The coupled-enzyme approach, on the other hand, is a more general and preferred approach. It utilizes the first enzyme for the desired biotransformation and a second enzyme and a second substrate for cofactor regeneration. Formate dehydrogenase[180,182] and glucose dehydrogenase (GDH)[186,187] are commonly used as the second enzyme for the regeneration of NAD(P)H. The coupled-enzyme concept has been successfully demonstrated using two isolated-enzymes[180,182,185-187], whole-cells co-expressing the both enzymes[181,183,184] and coupled permeabilized cells system[118]. Schrader and colleagues have also developed a recombinant *E. coli* expressing the $P450_{BM3}$ with an additional heterologous NADPH regeneration system comprising glucose facilitator and dehydrogenase. However, the product yield (based on cell dry weight) for this cofactor regeneration system did not show any improvement over the recombinant *E. coli* expressing the $P450_{BM3}$ alone.[160]

As described in Chapter 3, we have constructed a highly active *Escherichia coli* recombinant expressing the P450pyr system. Here, we are interested to increase the productivity of this recombinant strain by introducing a cofactor regeneration system *in vivo*. In this chapter, two cofactor regeneration systems utilizing either the glucose dehydrogenase[118] (GDH) from *Bacillus substilis* BGSC 1A1 or the engineered phosphite dehydrogenase 12x[33,123] (PTDH 12x) from *Pseudomonas stutzeri* were explored.

4.2 Results and Discussion

4.2.1 Construction of Recombinant *E. coli* Strains

We have constructed various *E. coli* strains expressing the four genes, including the P450pyr and ferredoxin (Fdx) genes from *Sphingomonas* sp. HXN200, the ferredoxin reductase (FdR) gene from *Mycobacterium* sp. HXN1500, and the cofactor regeneration component, GDH or PTDH 12x. These genes were inserted into various expression vectors in different combinations (Table 4.1). For example, P450pyr and GDH were inserted into the pRSFDuet, pCDFDuet, and pACYCDuet to construct pRSFDuet PG, pCDFDuet PG, and pACYCDuet PG, respectively. Similarly, Fdx and FdR were inserted into pETDuet to construct pETDuet FR. Similarly, Fdx and FdR were inserted into pETDuet to construct pETDuet FR. The constructed plasmids were then transformed into *E. coli* BL21(DE3) to obtain different strains with 2- and 3-plasmid systems (Figure 4.1).

Table 4.1. Various *E. coli* BL21(DE3) strains with 2- and 3-plasmid systems constructed in this study.

Strain	Host	Plasmids
Rp450	BL21(DE3)	pRSFDuet P450, pETDuet FR
Cp450	BL21(DE3)	pCDFDuet P450, pETDuet FR
Ap450	BL21(DE3)	pACYCDuet P450, pETDuet FR
A1	BL21(DE3)	pRSFDuet P450, pETDuet FR, pACYCDuet GDH
A2	BL21(DE3)	pRSFDuet P450, pETDuet FR, pACYCDuet 12x
A3	BL21(DE3)	pRSFDuet P450, pETDuet FR, pCDFDuet GDH
A4	BL21(DE3)	pRSFDuet P450, pETDuet FR, pCDFDuet 12x
Rgdh	BL21(DE3)	pRSFDuet PG, pETDuet FR
R12x	BL21(DE3)	pRSFDuet P12x, pETDuet FR
Cgdh	BL21(DE3)	pCDFDuet PG, pETDuet FR
C12x	BL21(DE3)	pCDFDuet P12x, pETDuet FR
Agdh	BL21(DE3)	pACYCDuet PG, pETDuet FR
A12x	BL21(DE3)	pACYCDuet P12x, pETDuet FR

FR: Fdx and FdR cloned into the same plasmid. PG: P450 and GDH cloned into the same plasmid. P12x: P450pyr and PTDH 12x cloned into the same plasmid.

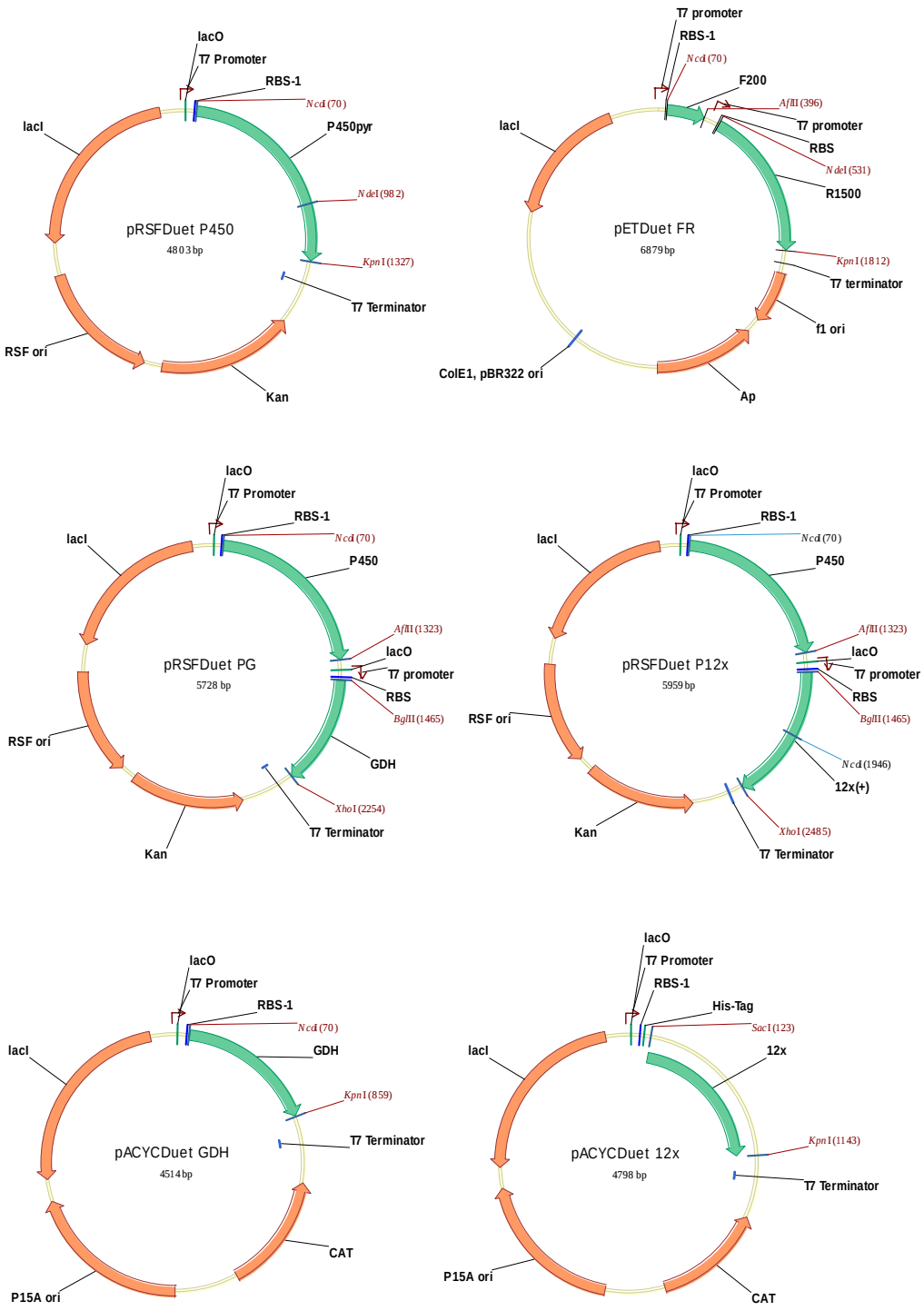


Figure 4.1. Selected examples of plasmid maps of pRSFDuet P450, pETDuet FR, pRSFDuet PG, pRSFDuet P12x, pACYCDuet GDH and pACYCDuet 12x. The green sections are the genes that were cloned into the plasmid.

4.2.2 Cell Culture and Protein Expression

As a quick initial test of activity, the *E. coli* BL21(DE3) strains were cultured and induced with 0.5 mM IPTG for protein expression. After 5 h of shaking at 250 rpm and 30 ºC, the cells were harvested and subjected to whole-cell assays as described in the Materials and Methods section. The strains with three plasmids grew the fastest but did not show any activity. When they were harvested after 5 h of induction, these cells were of the same color as the non-induced controls, rather than the dark pinkish color which would have been expected if P450pyr was expressed in the cell. When the cell-free extracts of these 3-plasmid strains were analyzed by SDS-PAGE, the P450pyr band was missing indicating that the P450pyr was hardly expressed at all. Hence, we switched our focus to the two-plasmid systems instead. The cell-free extracts of selected strains were analyzed by SDS-PAGE (Figure 4.2).

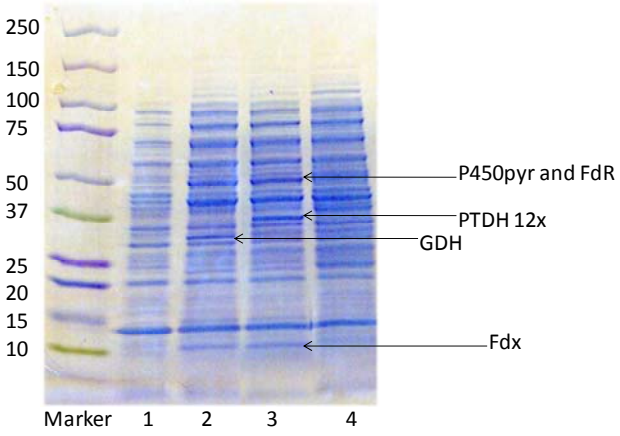
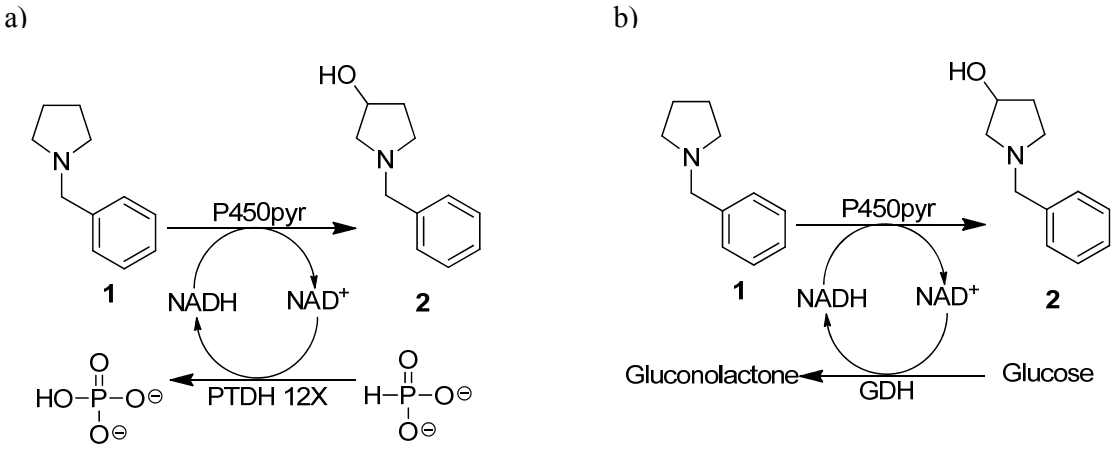


Figure 4.2. SDS-PAGE of non-induced R12x control (lane 1), Rgdh (lane 2), R12x (lane 3) and A2 (lane 4)

4.2.3 Biohydroxylation of *N*-Benzyl-pyrrolidine 1 with Recombinant *E. coli* Strains Expressing the P450pyr and Cofactor Regeneration System

The potential for increasing the productivity of *N*-benzyl-3-hydroxypyrrolidine **2** from the biohydroxylation of *N*-benzyl-pyrrolidine **1** was explored by introducing a cofactor regeneration system by utilizing either the GDH from *Bacillus substilis* BGSC 1A1 or the engineered PTDH 12x from *Pseudomonas stutzeri* (Scheme 4.1). The biohydroxylation of **1** to **2** is of great interest as these enantiomers are important pharmaceutical intermediates in the synthesis of a *k*-receptor agonist, an antibacterial agent, carbapenem antibiotics, and a 5-HT_{1Da} receptor agonist.[126] In the regeneration of NADH from NAD^+, the GDH converts glucose to gluconolactone whereas the PTDH 12x converts phosphite (HPO_3^{2-}) to phosphate (HPO_4^{2-}). For the GDH system, we investigated the biohydroxylation of three different strains, Rgdh, Cgdh, and Agdh (Table 4.1), in which the P450 and GDH genes were both expressed together on one of the following plasmids: pRSFDuet, pCDFDuet, or pACYCDuet. Similarly, for the PTDH 12x system, we used three different strains, R12x, C12x, and A12x. The activities and product concentration-time curves of these strains were compared to those of the parental strains, Rp450, Cp450, and Ap450. The whole-cell assay was carried out as described in Materials and Methods with samples taken at different time intervals.



Scheme 4.1. Biohydroxylation of substrate *N*-benzyl pyrrolidine **1** to its corresponding (*R*)- and (*S*)-*N*-benzyl-3-hydroxypyrrolidines **2** with NADH regeneration using (a) phosphite dehydrogenase (PTDH 12x) and (b) glucose dehydrogenase (GDH).

For the GDH cofactor regeneration system (Figure 4.3), the best strain was the Rgdh with a specific activity of 8.89 U/g cdw that was 1.6x higher than its parent strain Rp450 (without GDH gene) that has a specific activity of 5.48 U/g cdw (Table 4.2). The Rgdh strain produced up to 1.89 mM and 2.03 mM of **2** at 5 h and 26 h, respectively, compared to its parent strain Rp450 which produced only 1.44 mM and 1.75 mM of **2** at 5 h and 26 h, respectively. The Cgdh strain showed no improvement in its specific activity (or productivity) when compared to its parent Cp450 strain. On the other hand, the Agdh strain demonstrated a slight 1.2x increase in specific activity over its parent Ap450 strain, with 7.71 U/g cdw compared to 6.27 U/g cdw. At 26 h, it produced up to 2.34 mM of **2** compared to its parent Ap450 strain which produced 1.98 mM of **2**. Hence, the introduction of a GDH cofactor regeneration system into the P450pyr system increased its activity (and productivity), with the Rgdh strain showing the highest improvement over its parental strain.

Table 4.2. Specific activity for the biohydroxylation of **1** by various strains with the GDH and PTDH 12x systems. Activity was determined over the first 1h.

GDH system		PTDH 12x system	
Strain	**Specific Activity (U/g cdw)**	**Strain**	**Specific Activity U/g cdw**
Rp450	5.48	Rp450	3.48
Rgdh	8.89	R12x	8.67
Cp450	5.79	Cp450	4.28
Cgdh	5.02	C12x	3.49
Ap450	6.27	Ap450	5.22
Agdh	7.71	A12x	6.29

For the PTDH 12x cofactor regeneration system (Figure 4.4), the best strain was the R12x with a specific activity of 8.67 U/g cdw that was 2.5x higher than its parent strain Rp450 (without PTDH 12x gene) which has a specific activity of 3.48 U/g cdw. The R12x strain produced up to 1.96 mM and 2.28 mM of **2** at 5 h and 26 h, respectively, while its parent strain Rp450 produced only 1.00 mM and 1.70 mM of **2** at 5 h and 26 h, respectively. The C12x strain showed no improvement in its specific activity (or productivity) when compared to its parent Cp450 strain. On the other hand, the A12x strain demonstrated a slight 1.2x increase in specific activity over its parent Ap450 strain, with 6.29 U/g cdw compared to 5.22 U/g cdw. At 26h, it produced up to 2.20 mM of **2** compared to its parent Ap450 strain which produced 1.93 mM of **2**. Hence, the introduction of a PTDH 12x cofactor regeneration system into the P450pyr system increased its activity (and productivity), with the R12x strain showing the highest improvement over its parental strain.

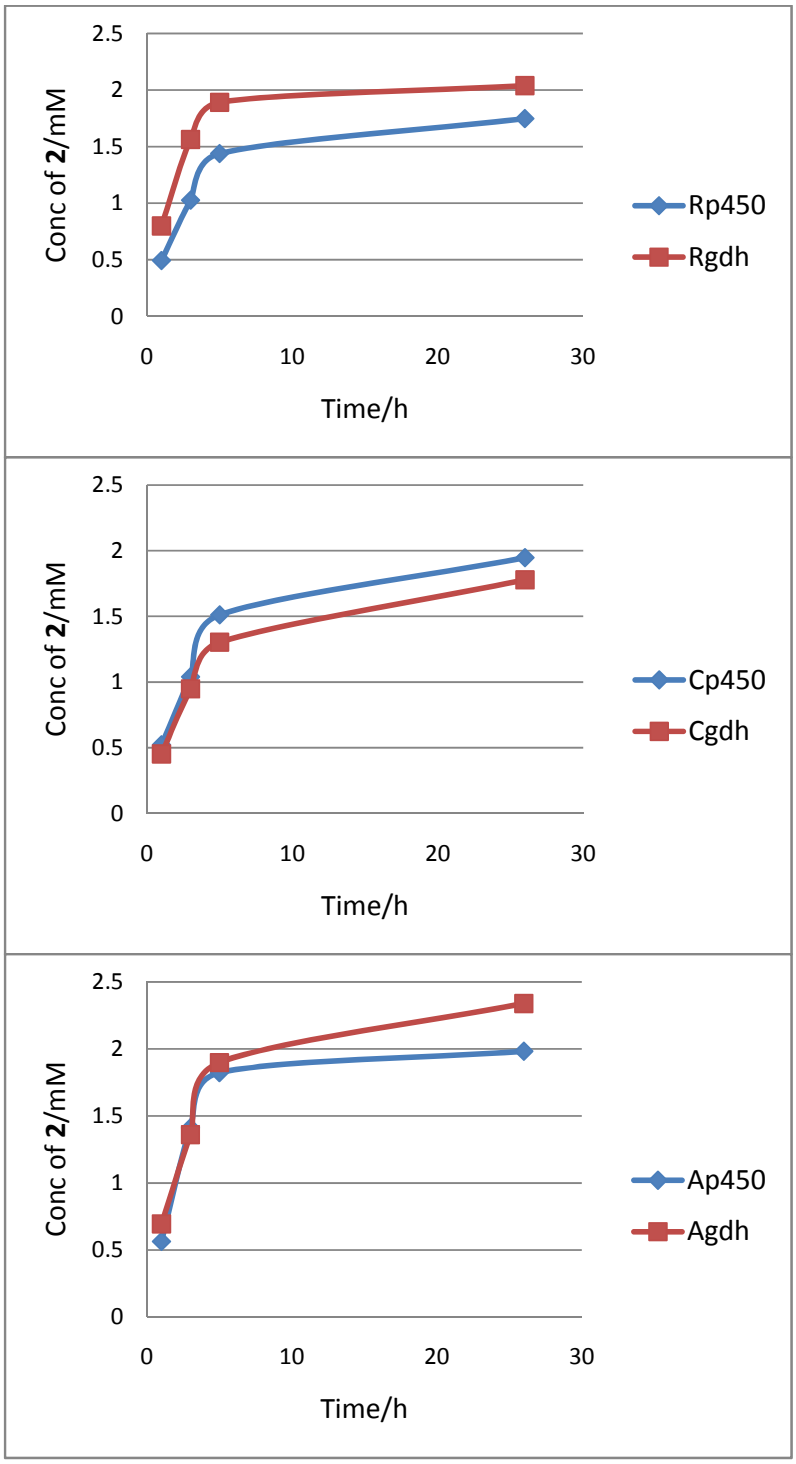


Figure 4.3. GDH cofactor regeneration system. Plot of product concentration (mM) versus time (h) for various strains. The whole-cell assay (10 mL total volume) was carried out in 125 mL shaking flasks with cell density of 1.5 g cdw/L, 8 % (w/v) glucose and 5 mM substrate **1** in 50 mM potassium phosphate buffer.

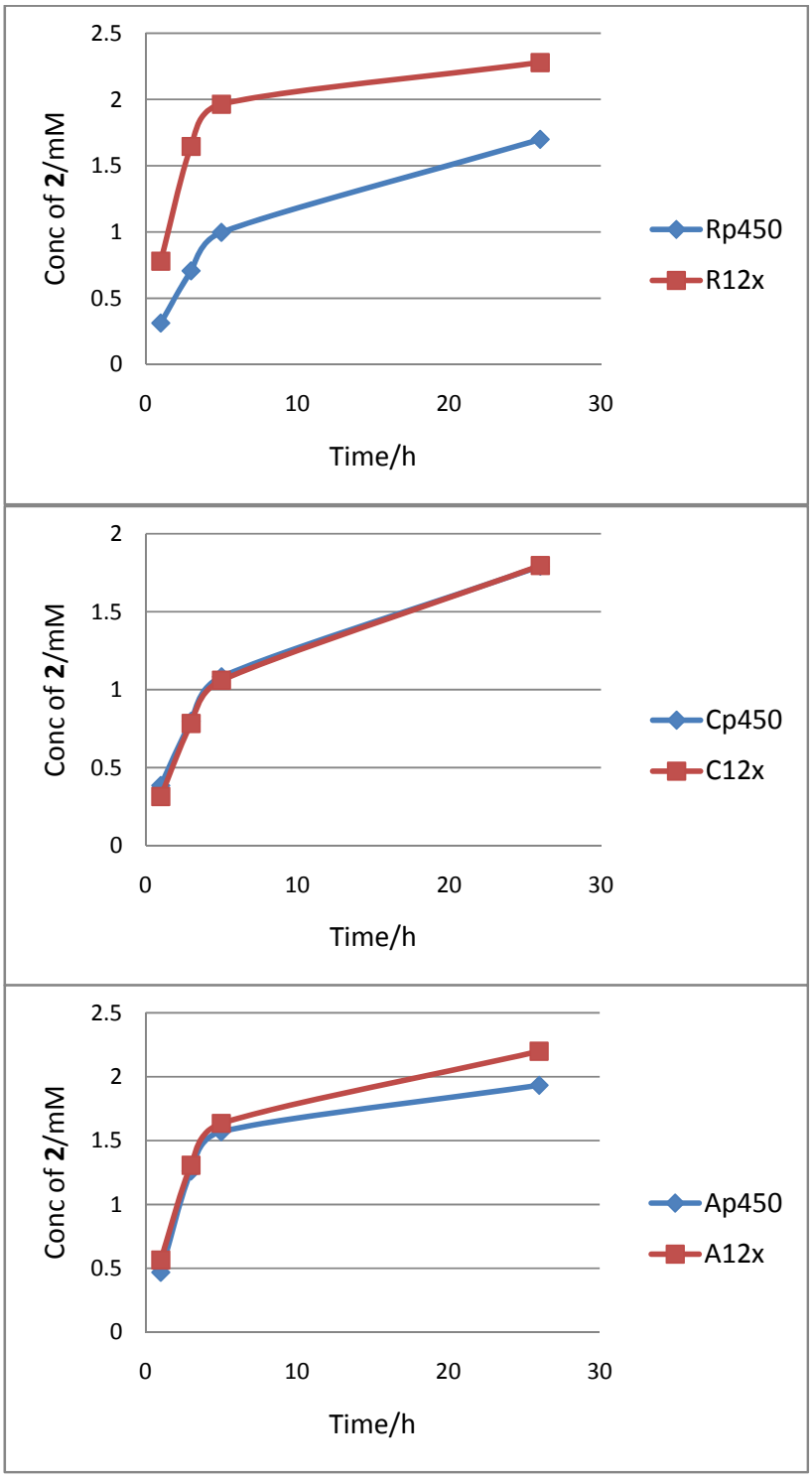


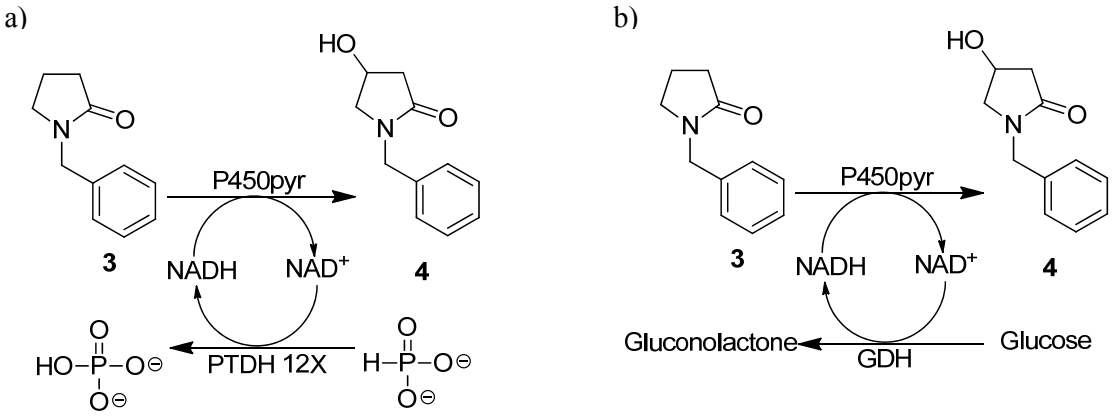
Figure 4.4. PTDH 12x cofactor regeneration system. Plot of product concentration (mM) versus time (h) for various strains. The whole-cell assay (10 mL total volume) was carried out in 125 mL shaking flasks with cell density of 1.5 g cdw/L, 2 % (w/v) glucose, 200 mM sodium phosphite and 5 mM substrate **1** in 50 mM potassium phosphate buffer.

4.2.4 Biohydroxylation of *N*-Benzyl-pyrrolidin-2-one 3 with Recombinant *E. coli* Strains Expressing the P450pyr and Cofactor Regeneration System

The potential for increasing the productivity of *N*-benzyl-4-hydroxypyrrolidin-2-one **4** from the biohydroxylation of *N*-benzyl-pyrrolidin-2-one **3** was explored by introducing a GDH cofactor regeneration system by utilizing either the GDH from *Bacillus substilis* BGSC 1A1 or the engineered PTDH 12x from *Pseudomonas stutzeri* (Scheme 4.2). This hydroxylation reaction is of great interest as it produces **4,** a very useful pharmaceutical intermediate, with an *ee* of > 99.9% (*S*). Similar to the previous section with substrate **1**, we investigated three different strains, Rgdh, Cgdh, and Agdh (Table 4.1) for the GDH system, and another three different strains, R12x, C12x, and A12x for the 12x system. The whole-cell assay was carried out as described in Materials and Methods with samples taken at different time intervals.

For the GDH cofactor regeneration system (Figure 4.5), the best strain was the Agdh with a specific activity of 5.02 U/g cdw that was 1.2x higher than its parent strain Ap450 (without GDH gene) which has a specific activity of 4.09 U/g cdw (Table 4.3). The Agdh strain produced up to 1.73 mM of **4** at 23 h compared to its parent strain Ap450 which produced only 1.32 mM of **4** at 23 h. The Cgdh strain showed no improvement in its specific activity (or productivity) when compared to its parent Cp450 strain. On the other hand, the Rgdh strain demonstrated a slight 1.1x increase in specific activity over its parent Rp450 strain, with 4.39 U/g cdw compared to 4.17 U/g cdw. At 23 h, it produced up to 1.43 mM of **4** compared to its parent Rp450 strain which produced 1.14 mM of **4**. Hence, the introduction of a GDH cofactor regeneration system into the P450pyr system increased its activity (and

productivity), with the Agdh strain showing the highest improvement over its parental strain.



Scheme 4.2. Biohydroxylation of substrate *N*-benzyl-pyrrolidin-2-one **3** to its corresponding (*S*)- and (*R*)-*N*-benzyl-4-hydroxypyrrolidin-2-ones **4** with NADH regeneration using (a) phosphite dehydrogenase (PTDH 12x) and (b) glucose dehydrogenase (GDH).

Table 4.3. Specific activity of various strains with the GDH and PTDH 12x systems. Activity was determined over the first 30 min.

	GDH system		**PTDH 12x system**
Strain	**Specific Activity (U/g cdw)**	**Strain**	**Specific Activity (U/g cdw)**
Rp450	4.17	Rp450	4.56
Rgdh	4.39	R12x	4.81
Cp450	4.64	Cp450	4.44
Cgdh	3.98	C12x	4.54
Ap450	4.09	Ap450	3.23
Agdh	5.02	A12x	5.91

For the PTDH 12x cofactor regeneration system (Figure 4.6), the best strain was the A12x with a specific activity of 5.91 U/g cdw that was 1.8x higher than its parent strain Ap450 (without PTDH 12x gene) which has a specific activity of 3.23 U/g cdw. The A12x strain produced up to 1.65 mM of **4** at 23 h compared to its parent strain Ap450 which produced only 1.36 mM of **4** at 23 h. The C12x strain showed no

improvement in its specific activity (or productivity) when compared to its parent Cp450 strain. On the other hand, the R12x strain demonstrated a very slight 1.1x increase in specific activity over its parent Ap450 strain, with 4.81 U/g cdw compared to 4.56 U/g cdw. Hence, the introduction of a PTDH 12x cofactor regeneration system into the P450pyr system increased its activity (and productivity), with the A12x strain showing the highest improvement over its parental strain.

We are interested to know the effect of increasing substrate **3** concentration on the whole-cell productivity curve for strains Ap450, Agdh and A12x (Figure 4.7). When the initial concentration of substrate **3** was increased to 10, 15 and 20 mM, the overall productivity fell by about 33% with each increase of 5 mM. The A12x strain has the highest productivity, followed by the Agdh. Nevertheless, the differences between all three strains became less obvious as initial substrate concentration was increased. This indicates that the substrate **3** has an inhibitory effect towards the enzyme activity. With a continuous supply of NADH regenerated by the PTDH 12x, the biohydroxylation reaction with the A12x strain continued up to the 10.5 h whereas the production of **4** with the Ap450 strain started to plateau off at 5 h. For the Agdh strain, the productivity is generally lower than that of the A12x strain but higher than that of the Ap450 strain. This can perhaps be attributed to the competition between our GDH system and the cell's natural metabolism for the glucose. Nevertheless, at initial substrate concentration of 10 mM, the Agdh strain continuously catalyzed the biohydroxylation of substrate **3** to product **4** even at the 26^{th} hour.

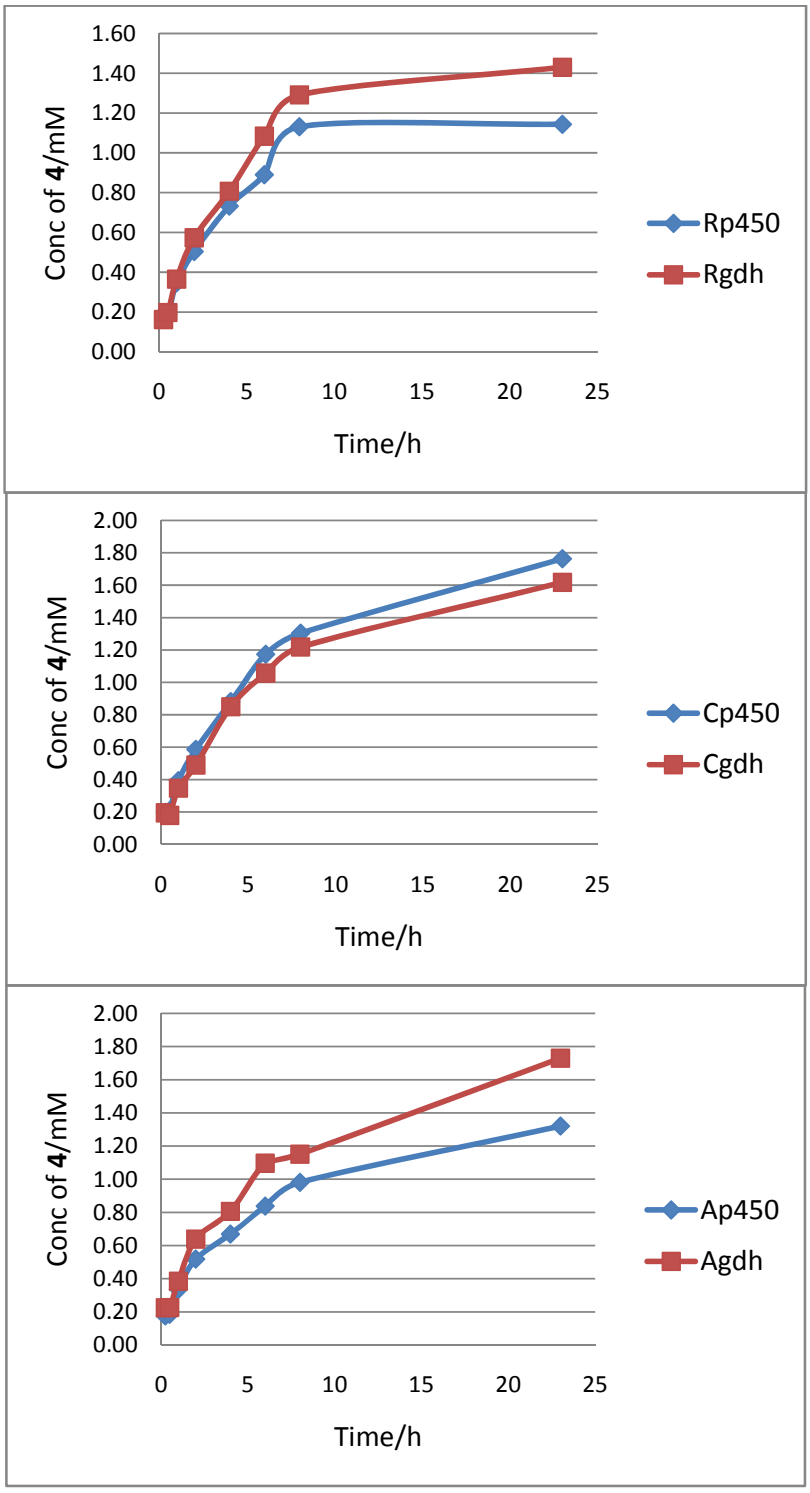


Figure 4.5. GDH cofactor regeneration system. Plot of product concentration (mM) versus time (h) for various strains. The whole-cell assay (10 mL total volume) was carried out in 125 mL shaking flasks with cell density of 1.5 g cdw/L, 2 % (w/v) glucose and 5 mM substrate **3** in 50 mM potassium phosphate buffer.

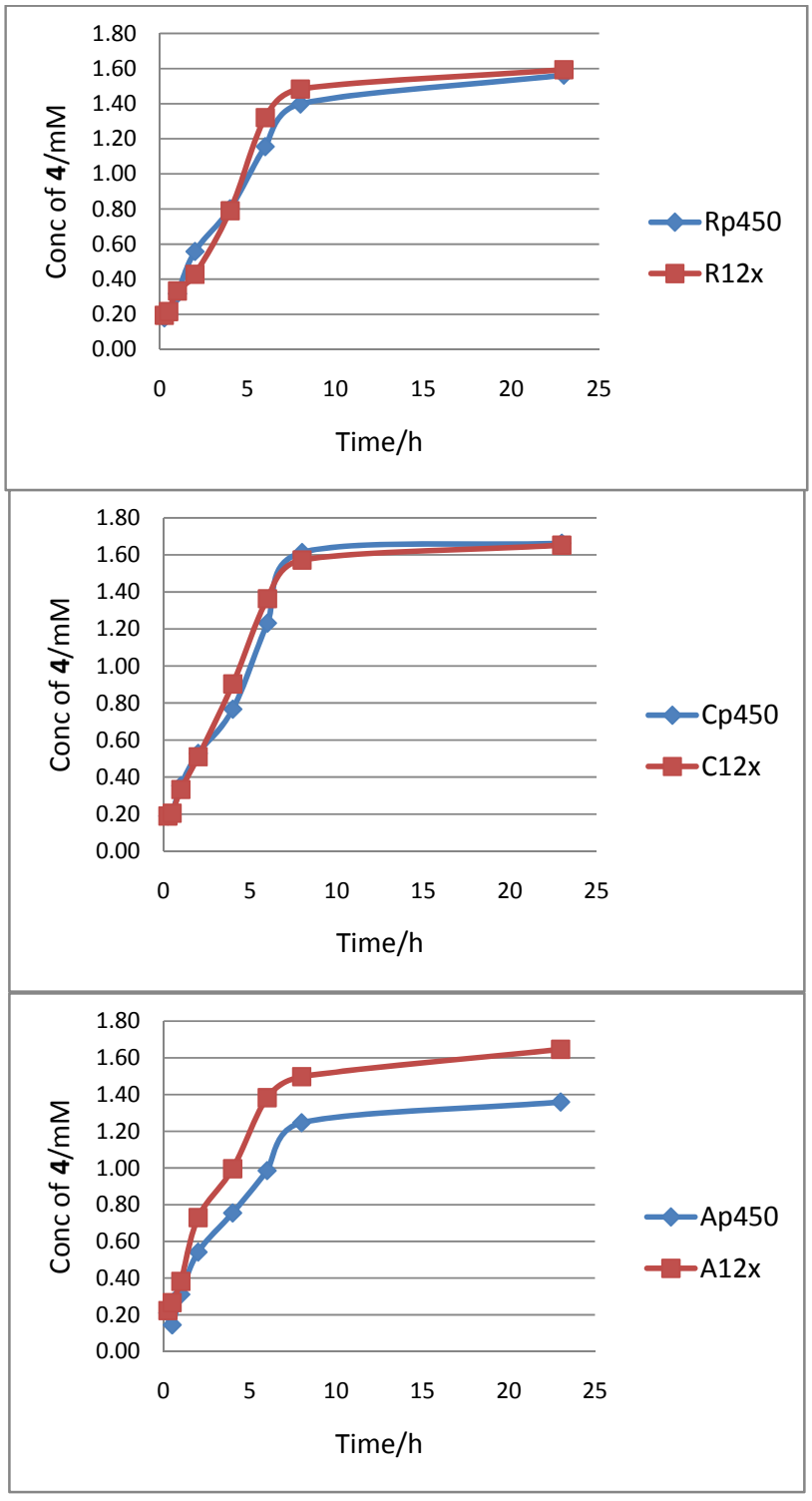


Figure 4.6. PTDH 12x cofactor regeneration system. Plot of product concentration (mM) versus time (h) for various strains. The whole-cell assay (10 mL total volume) was carried out in 125 mL shaking flasks with cell density of 1.5 g cdw/L, 2 % (w/v) glucose, 200 mM sodium phosphite and 5 mM substrate **3** in 50 mM potassium phosphate buffer.

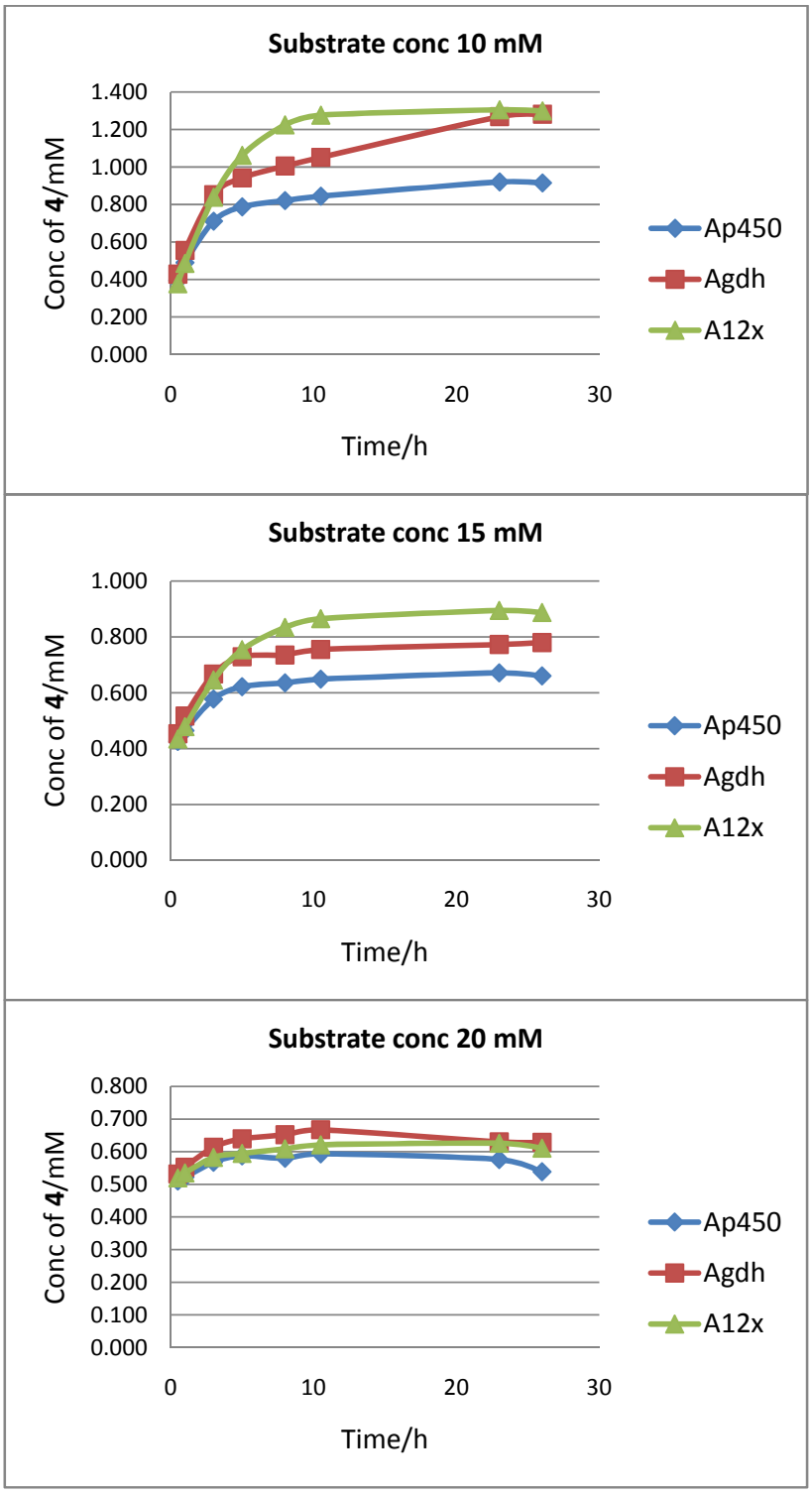


Figure 4.7. Comparison of strain productivity (for Ap450, Agdh and A12x) with different starting concentration of substrate **3** (10, 15 and 20 mM). Biotransformation was performed with cell density of 1.5 g cdw/L.

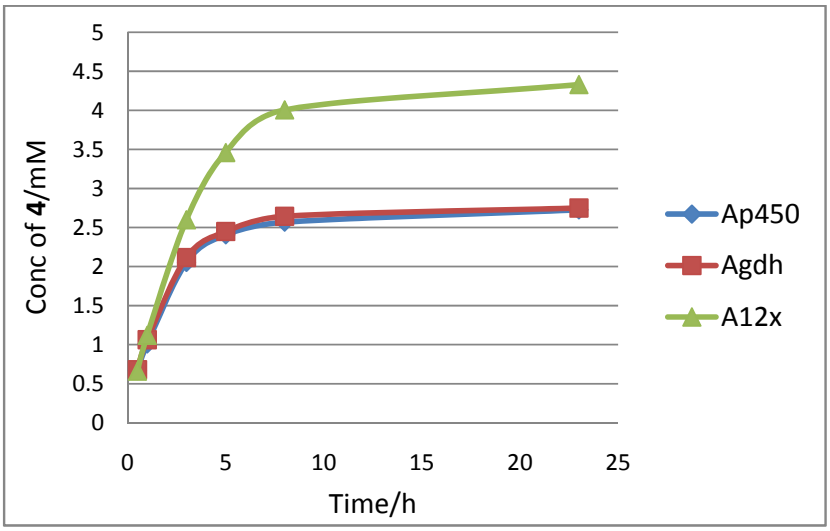


Figure 4.8. Plot of product concentration (mM) versus time (h). Biotransformation was performed with cell density of 5 g cdw/L, 2 % (w/v) glucose, 200 mM sodium phosphite (for PTDH 12x system) and 10 mM substrate **3** in 50 mM potassium phosphate buffer.

When a higher cell density was used in the biotransformation with an initial substrate **3** concentration of 10 mM (Figure 4.8), the A12x strain produced up to 4.33 mM of **4** at 23 h compared to the parent strain Ap450 of just 2.72 mM of **4,** a 1.6x improvement in productivity. Unexpectedly, the product concentration-time curve for the Agdh strain did not show any difference in productivity when compared to the parent strain Ap450. This is perhaps related to the competition for glucose by the newly introduced GDH system and the cell's natural GDH system.

Although *E. coli* BL21(DE3) has its own internal cofactor recycling system, the incorporation of an additional cofactor recycling system *in vivo* increased the availability of NADH in the system, thus enabling the P450pyr to be utilized more efficiently. Also, the product formation period was also prolonged thus increasing the productivity of the P450pyr.

Productivity improvement varied when different expression vectors were used due to the different copy number of plasmids in the system, resulting in different ratios of protein expressed within the system. One could also expect the expression level of the proteins to be different depending on the order it was cloned.

The addition of 200 mM of sodium phosphate into the PTDH 12x system has also been found to affect the specific activity for Rp450, Cp450 and Ap450 (Table 4.2). This can be attributed to the high salt concentration in the solution, thus affecting the ionic strength in the reaction mixture.

The conversion of glucose to gluconolactone resulted in the lowering of the pH of the whole-cell reaction mixture which would inadvertently affect the optimum activity of the enzymes. Hence, the pH would have to be checked at every 3-4 h interval and readjusted back to pH 7.5. On the other hand, for the PTDH 12x system, phosphate accumulation is not a problem as P450pyr and PTDH 12x are not inhibited by phosphate, which can also be used as to buffer the assay, and phosphate can be easily removed downstream if necessary by ionic filtration or precipitation as a calcium salt.

4.3 Conclusion and Outlook

The incorporation of a GDH or PTDH 12x cofactor regeneration system together with the P450pyr system in a recombinant *E. coli* has resulted in an increase in activity and productivity, i.e. the product formation period was prolonged. Using substrate **1**, the R12x strain demonstrated 2.5x higher activity than its parent strain

Rp450, whereas the Rgdh strain gave a 1.6x higher activity than the parent strain. On the other hand, when utilizing substrate **3**, the A12x strain demonstrated 1.8x higher activity than its parent strain Ap450, whereas the Agdh strain demonstrated only a 1.2x increase compared to its parent strain. When a cell density of 5 g cdw/L and 10 mM initial substrate **3** were used, up to 4.33 mM of **4** (*ee* > 99.9% (*S*)) was produced by A12x at 23 h compared to the parent strain Ap450 with just 2.72 mM of **4**. Further optimization work would be necessary to increase the productivity level of this system. This method could also be applied to other microbial biohydroxylation reactions.

4.4 Materials and Methods

Materials

Escherichia coli BL21(DE3) and the expression plasmids were purchased from Novagen (Madison, WI). PCR grade dNTPs were obtained from Roche Applied Sciences (Indianapolis, IN). Restriction enzymes, Phusion High-Fidelity DNA polymerase, T4 DNA ligase and their corresponding buffers were purchased from New England Biolabs (NEB) (Beverly, MA). D-glucose was purchased from ThermoFisher (Pittsburgh, PA). Ampicillin, kanamycin, isopropyl β-D-thiogalactopyranoside (IPTG), nitro blue tetrazolium (NBT), phenazine methosulfate (PMS), NAD^+, and $NADP^+$ were purchased from Sigma (St. Louis, MO). The Ni-NTA agarose, QIAprep spin plasmid mini-prep kit, QIAEX II gel purification kit, and QIAquick PCR purification kit were purchased from Qiagen (Valencia, CA). Various oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA). SDS-PAGE gels, buffers and protein size markers were purchased from Bio-Rad (Hercules, CA). All other required chemical reagents and solvents were purchased from Acros Organics (Morris Plains, NJ) and Sigma-Aldrich (St. Louis, MO).

Strain Construction

The P450pyr was cloned into pRSFDuet, and the Fdx and FdR were cloned into the pETDuet as described in Chapter 3. To construct the plasmids listed out on Table 4.1, the P450pyr and GDH or PTDH 12x genes were amplified by PCR using an MJ Research PTC-200 thermal cycler with the following program:

PCR program:

1 = 98 °C for 0:30
2 = 98 °C for 0:10
3 = 55-65 °C for 0:45
4 = 72 °C for 0:30 to 1:00
5 = Goto 2, 24 times
6 = 72 °C for 10:00
7 = 4 °C for ever
8 = End

The PCR amplified fragments were then digested with their respective enzymes and cloned into separate plasmids (Table 4.4). For plasmids with two genes, the P450pyr was ligated into the plasmid first, followed by the second gene, GDH or PTDH 12x. The resulting plasmids from the ligations were separately transformed into chemical competent *E. coli* DH5α cells which were then plated on Luria-Bertani (LB) agar plates with appropriate antibiotics. Finally, the purified plasmids were transformed into electrocompetent *E. coli* BL21(DE3) cells and plated on LB agar plates containing the appropriate antibiotics.

Table 4.4. Construction of different plasmids. Primers and restriction sites used are shown below. The same genes were ligated into different vectors as listed below.

Gene to be amplified	**Primers (Restriction sites underlined)**	**Cloned into …**
P450pyr Template: pCom8-PA7F200R1500[131]	PF200fwd NcoI: TTA ACT ACT CCA TGG AAC ATA CAG GAC AAA GCG CGG P450 rev AflII: TTA ACT ACT CTT AAG CTA CGC GTG GAC GCG AAC	MCS I of pRSFDuet, pCDFDuet and pACYCDuet
GDH Template: pUC18 GDH (Obtained from Prof Li Zhi's lab strain collection)	GDH fwd BglII: TTA ACT ACT AGA TCT ATG TAT CCG GAT TTA AAA GGA AAA GTC GTC GDH rev XhoI: TTA ACT ACT CTC GAG TTA ACC GCG GCC TGC C	MCS II of pRSFDuet, pCDFDuet and pACYCDuet

Table 4.4 (cont.)

	GDH fwd NcoI: TTA ACT ACT CCA TGG AT CCG GAT TTA AAA GGA AAA GTC GTC GDH rev KpnI: TTA ACT ACT GGT ACC TTA ACC GCG GCC TGC C	pACYCDuet
PTDH 12x Template: pET15b 12x (Obtained from Prof Zhao's lab strain collection)	12X F BglII: TTA ACT ACT AGA TCT ATG CTG CCG AAA CTC GTT ATA ACT C 12X R XhoI: TTA ACT ACT CTC GAG TTA TCA GTC TGC GGC AGG ATT G	MCS II of pRSFDuet, pCDFDuet and pACYCDuet
	12X F SacI: TTA ACT ACT GAG CTC ATG CTG CCG AAA CTC GTT ATA ACT C 12X R KpnI: TTA ACT ACT GGT ACC TTA TCA GTC TGC GGC AGG ATT G	pACYCDuet

Cell Culture and Protein Expression

The recombinant *E. coli* strain was streaked on an LB agar plate with appropriate antibiotics and incubated overnight at 37 ºC. A single colony from the plate was inoculated into 5 mL of LB media with appropriate antibiotics and was grown overnight at 37 ºC and 250rpm. Next, 25 mL of TB media with the appropriate antibiotics was inoculated with 1% (v/v) of the overnight culture and grown to mid-log phase (OD_{600} of ~0.6). The gene expression was induced by adding 0.5 mM IPTG and 0.5 mM δ-aminolevulinic acid (δ-ALA), and the cells were continuously shaken at 250 rpm at 30 ºC for 5 h. The cells were then harvested by centrifugation and resuspended in 50 mM potassium phosphate buffer (pH 7.5) to a cell density of 10 g cdw/L.

The cell suspension (10 g cdw/L) was passed twice through a homogenizer at 20 kpsi. The cell debris was removed by centrifugation at 13,000 rpm at 4 ºC for 20 min. The supernatant was diluted and subjected to analysis with SDS-PAGE.

Whole-Cell Assay

For GDH-based systems, a whole-cell assay was prepared with 1.5 g cdw/L of cell suspension, 2 to 8 % (w/v) glucose, 5 to 20 mM substrate **1** or **3** in 50 mM potassium phosphate buffer (pH 7.5) (Total volume 10 mL).

For PTDH 12x-based systems, a whole-cell assay was prepared with 1.5 to 5 g cdw/L of cell suspension, 0.2 M sodium phosphite, 2 % (w/v) glucose, 5 to 20 mM substrate **1** or **3** in 50 mM potassium phosphate buffer (pH 7.5) (Total volume 10 mL).

The mixtures were shaken at 30 ºC and 250 rpm. At different time points, 200 μL samples were taken and centrifuged for 2 min at 13,000 rpm. The supernatant from these samples were then subjected to HPLC analysis.

Analytical Method

The concentration of *N*-benzyl-3-hydroxypyrrolidine **2** was analyzed by Shimadzu Prominence HPLC using a Agilent Eclipse Plus C18 (3.5 μm) column (4.6 x 150 mm), UV detection at 210 nm, flow rate of 1 mL/min; eluent A, acetonitrile, and eluent B, 10 mM potassium phosphate buffer (pH 7.0). Retention times (min): t_R of **1**, 10.9; t_R of **2**, 6.8, A/B 20:80.

The concentration of *N*-benzyl-4-hydroxypyrrolidin-2-one **4** was analyzed by Shimadzu Prominence HPLC using a Hypersil BDS-C18 column (4 x 125 mm), UV

detection at 210 nm, flow rate of 1 mL/min; eluent A, acetonitrile, and eluent B, water. Retention times (min): t_R of **1**, 7.8; t_R of **2**, 2.8, A/B 20:80.

The *ee* of the **2** and **4** were analyzed by using the following conditions: Chiral column (250 m × 4.6 mm). UV detection at 210 and 254 nm, flow rate at 1.0 mL/min (unless stated otherwise); eluent A, *n*-hexane; eluent B, isopropanol. Retention times (min): t_R of (*R*)-**2**, 26.1; t_R of (*S*)-**2**, 43.5, column, chiralcel OB-H (Daicel), A/B 98/2; t_R of (*S*)-**4**, 20.3; t_R of (*R*)-**4**, 30.5, column, chiralpak AS, A/B 4/1.

Chapter 5 : Further Characterization of P450pyr and Related Mutants

5.1 Introduction

Cytochrome P450 is a ubiquitous monooxygenase superfamily that can catalyze the insertion of an oxygen atom into C-H or C=C bonds. Its ability to catalyze regio- and stereoselective hydroxylation of non-activated carbon atoms plays a key role in biosynthetic pathways as well as in chemical synthesis of important intermediates for fine chemicals and pharmaceuticals. P450 enzymes also play a prominent role in the synthesis and metabolism of secondary metabolites such as prostaglandins,[188] leucotrienes and thromboxanes,[189] steroid hormones,[190] insect and plant hormones,[191] or colors and odors in plants.[192] They are involved in the oxidative metabolism of pharmaceuticals and xenobiotic compounds in humans as well as bio-degradation of environmental hazardous chemicals. In mammalian liver, P450 monooxygenases function as a biological defense system.[193]

Bacterial P450s are more promising than those from plants and animals as they are mostly soluble, not membrane-associated, show higher reaction rates compared to P450s from other sources and exhibit a relatively high stability. To date, many of the bacteria P450s have been well characterized, such as P450BM-3 from *Bacillus megaterium,*[194] CYP102A2 from *Bacillus subtilis*[195] and CYP505 from *Fusarium oxysporum,*[196] P450cam from *Pseudomonas putida*[197] and P450nor from *F. oxysporum*[198] among many others. Recently, P450cam, P450terp and P450BM-3 were shown to catalyze the oxidation of alternative substrates, such as the oxidation of substituted thioanisoles to sulfoxide and the epoxidation of substituted styrenes,[199]

with rather poor selectivity. The well-characterized camphor hydroxylase P450cam from *Pseudomonas putida* has been engineered to convert other substrates, including halogenated aromatic compounds such as alkylchlorobenzene.[200,201] Another highly characterized and engineered P450 monooxygenase is P450BM-3 - a catalytically self-sufficient monooxygenase, which contains a heme domain and a flavin reductase domain on a single polypeptide chain.[202,203] This self-sufficient system facilitates *in vitro* applications because the auxiliary electron transport protein partners are not required to be added separately for function. P450BM-3 has also been extensively evolved to a P450 propane monooxygenase ($P450_{PMO}$) having 20 heme domain substitutions compared to WT P450BM-3.[44]

We are interested in the P450pyr from *Sphingomonas* sp. HXN-200 which belongs to a class I P450 monooxygenase and has been found to catalyze the hydroxylation of a broad range of substrates, such as *n*-alkanes, cyclic alkanes, *N*-heterocycles, spiro-oxazolidines, and benzoxazole derivatives, with highly activity, high regioselectivity, and good to excellent enantioselectivity.[10,126-130] In this chapter, we will focus on the partial characterization of cytochrome P450pyr and its enantioselective mutants that were previously engineered by stepwise site saturation mutagenesis. The WT P450pyr and its mutants 1AF4, 1AF4A and 11BB12 together with their auxiliary electron transport proteins, ferredoxin and ferredoxin reductase has also been successfully cloned and expressed in *E. coli* BL21(DE3).

5.2 Results

5.2.1 Cloning, Expression, and Purification of WT P450pyr and Its Mutants

The WT P450pyr and ferredoxin (Fdx) gene from *Sphingomonas* sp. HXN200 and the ferredoxin reductase (FdR) gene from *Mycobacterium* sp. HXN1500 were amplified from a plasmid pCom8-PA7F200R1500[131]. Meanwhile, the 1AF4, 1AF4A and 11BB12 mutants were amplified from the plasmid pRSFDuet containing the mutant P450s which were identified in the earlier rounds of directed evolution (Chapter 3). Next, each PCR-amplified gene was digested with its respective enzymes and cloned into the pET28a vector such that each gene would have an N-terminal His_6-tag. Subsequently, the resulting plasmids were chemically transformed into chemical competent *E. coli* DH5α cells. Finally, the purified plasmids containing the correct gene was electroporated into electrocompetent *E. coli* BL21(DE3).

The constructed recombinant strains were each cultured separately in TB medium containing 50 μg/mL of kanamycin at 37 ºC. The gene expression was induced by adding 0.5 mM IPTG during the mid to late exponential phase, and the cells were continuously shaken at 250 rpm at 23 ºC for ~16-18 h. Standard enzyme purification protocol used has been described in the Materials and Methods section. SDS-PAGE analysis clearly revealed a thick band with the expected size of approximately 48 kDa for P450pyr, 47 kDa for FdR and 12 kDa for Fdx (Figure 5.1). The purified protein is estimated to have a purity of >90% based on the overloaded SDS-PAGE gel. The concentration of the purified protein was quantified using the Bradford protein assay. The CO difference spectrum was also used to quantify the active P450 concentration.

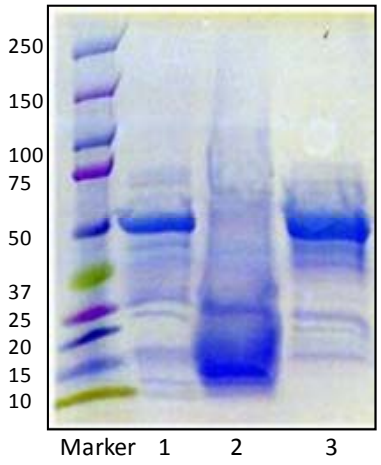


Figure 5.1. SDS-PAGE of purified proteins. P450pyr (lane 1), Fdx (lane 2) and FdR (lane 3).

5.2.2 *In vitro* Kinetic Analysis

The steady state kinetic parameters for WT P450pyr, 1AF4, 1AF4A and 11BB12 were determined from initial velocity measurements of the formation of product *N*-benzyl-3-hydroxypyrrolidines **2** from substrate *N*-benzyl pyrrolidine **1.** The P450pyr and its mutants were reconstituted with the electron transport proteins Fdx and FdR. Initial activity assay indicated that increasing the Fdx concentration from 1 to 10 μM while keeping the P450pyr and FdR concentration constant at 1 μM increases the specific activity of the P450pyr (Table 5.1). On the other hand, when the FdR concentration is increased from 1 to 10 μM, the specific activity of the P450 decreases. When FdR is removed from the *in vitro* system, no activity was observed, confirming the fact that FdR is required for P450pyr activity. Further increase of the Fdx concentration beyond 10 μM showed no further increase in the P450 specific activity. Thus, we concluded that the optimum P450:Fdx:FdR is 1:10:1.

Table 5.1. Optimizing the ratio of P450:Fdx:FdR. Specific activity increases with increasing Fdx but decreases with increasing FdR. The reaction was conducted in the presence of the relevant enzymes in 50 mM Tris HCl buffer, pH 7.5, containing 1 mM substrate **1** and 1 mM NADH. Total volume was 500 µL. Product **2** formation was measured at 30 minutes.

	Concentration / µM			
Sample	P450pyr	Fdx	FdR	Product 2 formation
1	1	1	1	10.76
2	1	5	1	44.20
3	1	10	1	81.38
4	1	20	1	81.09
5	1	30	1	81.09
6	1	40	1	62.33
7	1	10	0	0
8	1	5	1	47.04
9	1	5	5	19.53
10	1	5	10	8.27

Using this ratio, we proceeded to determine the saturating NADH concentration (P450 concentration is fixed at 0.5 µM) using different concentrations of NADH from 0.1 mM to 1 mM. The specific activity of the P450pyr reached its maximum at 0.2 mM NADH. Thus, we decided to use 1 mM NADH per 0.5 µM of P450, i.e. a five-fold oversaturation which should be more than sufficient.

Table 5.2. Steady state kinetic parameters of WT P450pyr and its mutants 1AF4, 1AF4A and 11BB12. The values for k_{cat}, K_M and k_{cat}/K_M are based on the average of three separate kinetic assays. The kinetic assay for both 11BB12 and 1AF4 was based on an enzyme concentration of 0.5 µM whereas the 1AF4A kinetic assay was based on 1 µM.

Sample	k_{cat} (min^{-1})	K_M (mM)	k_{cat}/K_M
WT P450	10.23	1.78	6.08
11BB12	2.42	4.27	0.58
1AF4	2.27	6.90	0.33
1AF4A	1.11	9.11	0.13

Using the optimized conditions as determined previously, we ran the kinetic assays for WT P450, 1AF4, 1AF4A and 11BB12 (Table 5.2). When compared to the

WT P450pyr, the k_{cat} of the mutants was reduced by more than four-fold whereas the K_M was increased by more than two-fold. Among the mutants, 11BB12 has the highest k_{cat} of 2.42 min^{-1} and the lowest K_M value of 4.27 (based on 0.5 μM of enzyme). 1AF4 shared a similar k_{cat} value to 11BB12 although the K_M value is 1.6 fold higher at 6.9 mM. 1AF4A also demonstrated the lowest k_{cat} of only 1.11 min^{-1} and a high K_M of 9.11 (based on 1 μM of enzyme).

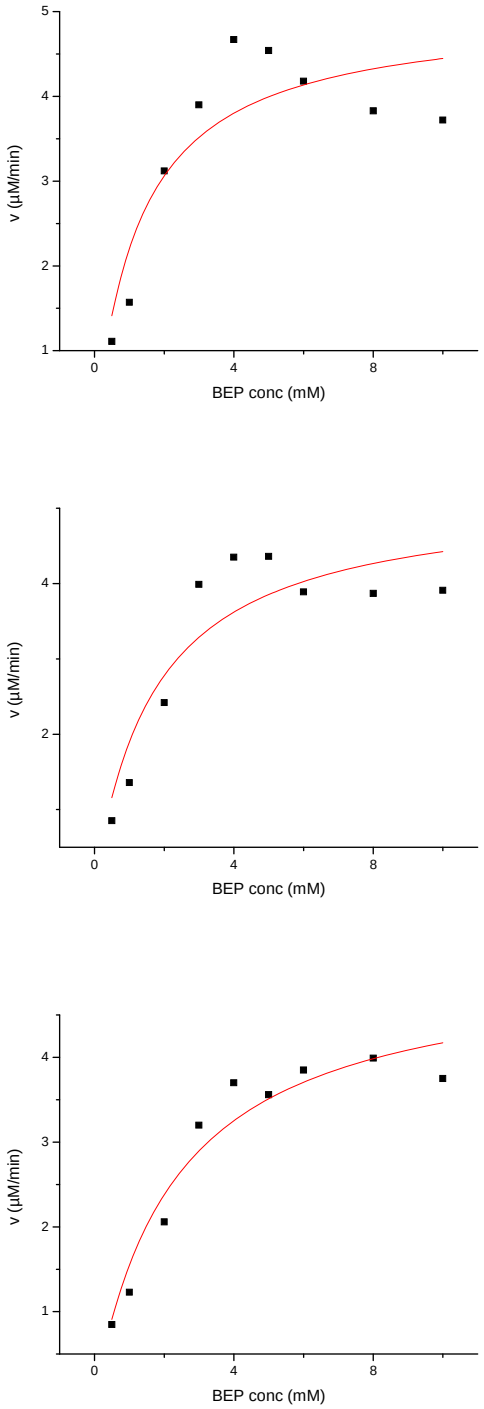


Figure 5.2. Michaelis-Menten plot for WT P450pyr.

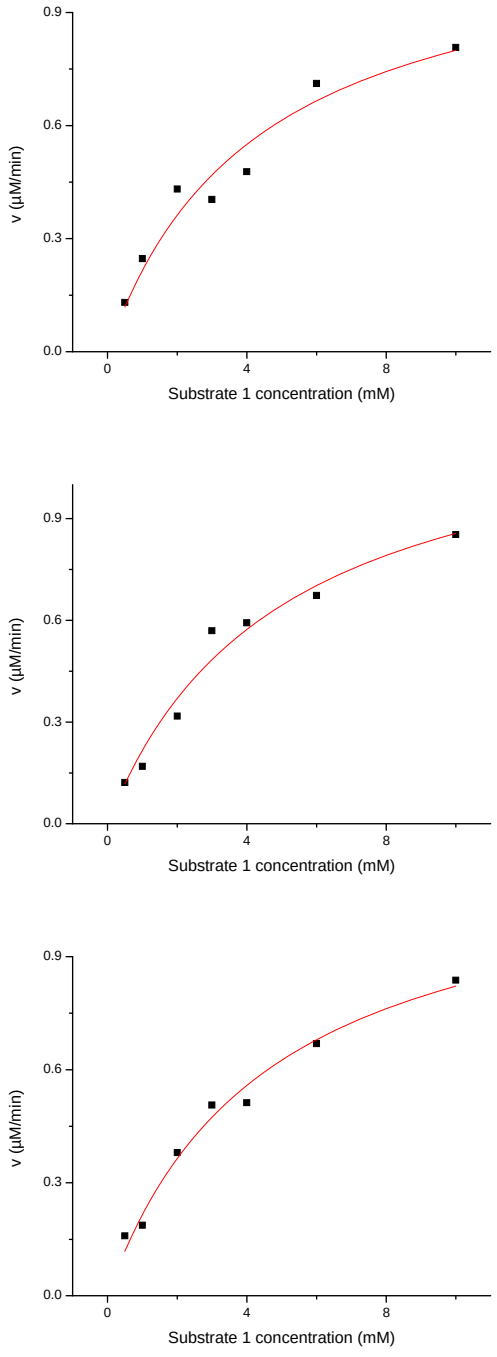


Figure 5.3. Michaelis-Menten plot for 11BB12. The maximum substrate concentration used was capped at 10 mM as the solubility of the substrate in aqueous solution is noticeably reduced beyond 15 mM.

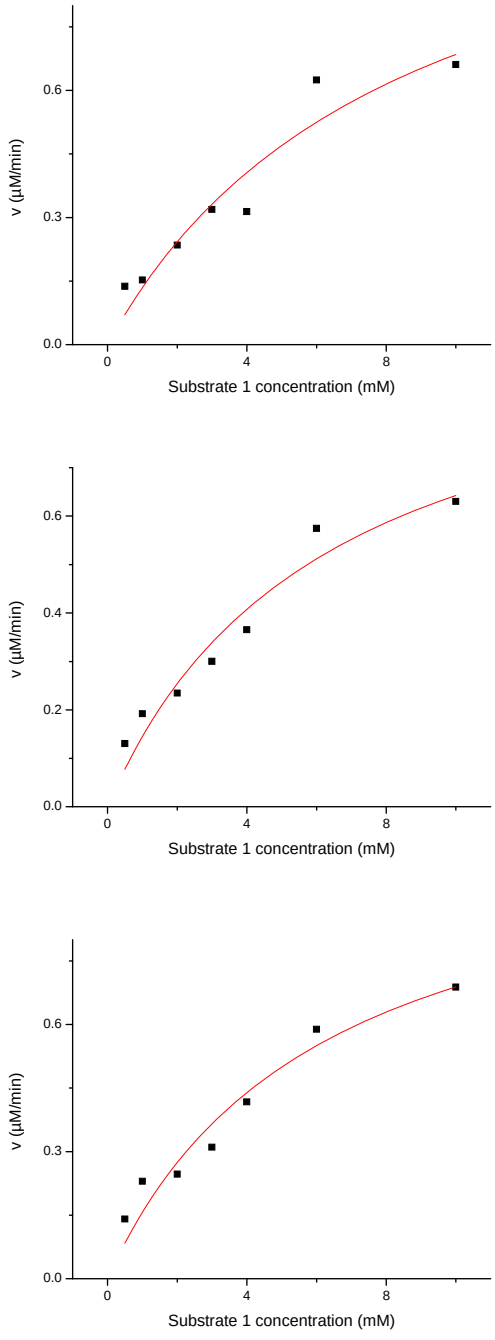


Figure 5.4. Michaelis-Menten plot for 1AF4. The maximum substrate concentration used was capped at 10 mM as the solubility of the substrate in aqueous solution is noticeably reduced beyond 15 mM.

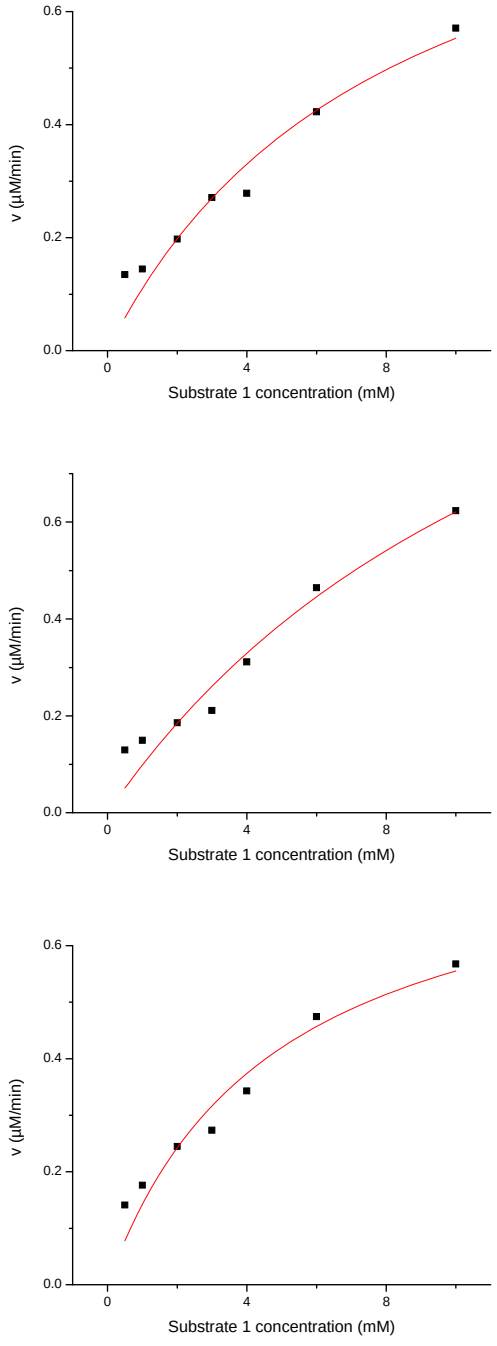


Figure 5.5. Michaelis-Menten plot for 1AF4A. The maximum substrate concentration used was capped at 10 mM as the solubility of the substrate is noticeably reduced beyond 15 mM.

5.2.3 Biohydroxylation of Mutant P450s with Different Substrates

WT P450pyr has previously been shown to have a wide substrate range, including piperidines, pyrrolidinones and piperidinones.[127-130] Hence, we are interested in investigating the changes in product *ee* when 1AF4 and 1AF4A mutants are used to catalyze the hydroxylations of a range of different substrates using whole-cell assays. In this experiment, we have biohydroxylated following substrates to their corresponding alcohols: *N*-benzyl pyrrolidine **1** (our original substrate), *N*-benzyloxycarbonyl-pyrrolidine **3**, *N*-Boc-pyrrolidine **5**, *N*-benzyl pyrrolidinone **7**, *N*-Boc- pyrrolidinone **9**, *N*-benzyl piperidinone **11**, and *N*-Boc-piperidinone **13** (Figure 5.6).

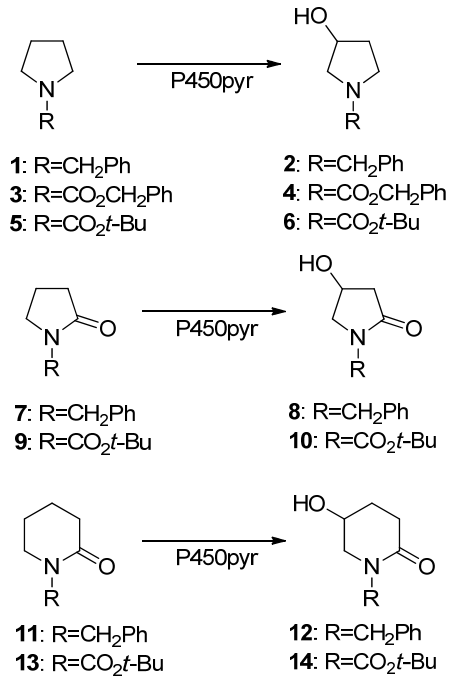


Figure 5.6. Hydroxylation scheme of various substrates to their respective products.

Table 5.3. Product *ee* of various substrates.

Substrates	WT *ee* (%)	1AF4 *ee* (%)	1AF4A *ee* (%)
1	43 (*S*)	42 (*R*)	83 (*R*)
3	75 (*R*)	78 (*R*)	Formation of a side product which overlaps with the enantiomeric peaks
5	4 (*R*)	8 (*R*)	21 (*R*)
7	> 99.9 (*S*)	99 (*S*)	99 (*S*)
9	90-92 (*S*)	90 (*S*)	90 (*S*)
11	55 (*S*)	60 (S)	44 (R)
13	28 (*R*)	9 (*R*)	73 (*S*)

From a previous directed evolution work, we found that both 1AF4 and 1AF4A can invert the WT P450 enantioselectivity towards substrate **1**. Here, we found that both 1AF4 and 1AF4A can also change the enantioselectivity towards substrate **11** and **13** (Table 5.3). In fact, 1AF4A could invert the *ee* of substrate **11** and **13** (Figure 5.7 and 5.8). There was hardly a change in enantioselectivity for substrates **3**, **7** and **9**. Although the enantioselectivity was slightly improved towards substrate **5**, it is still poor.

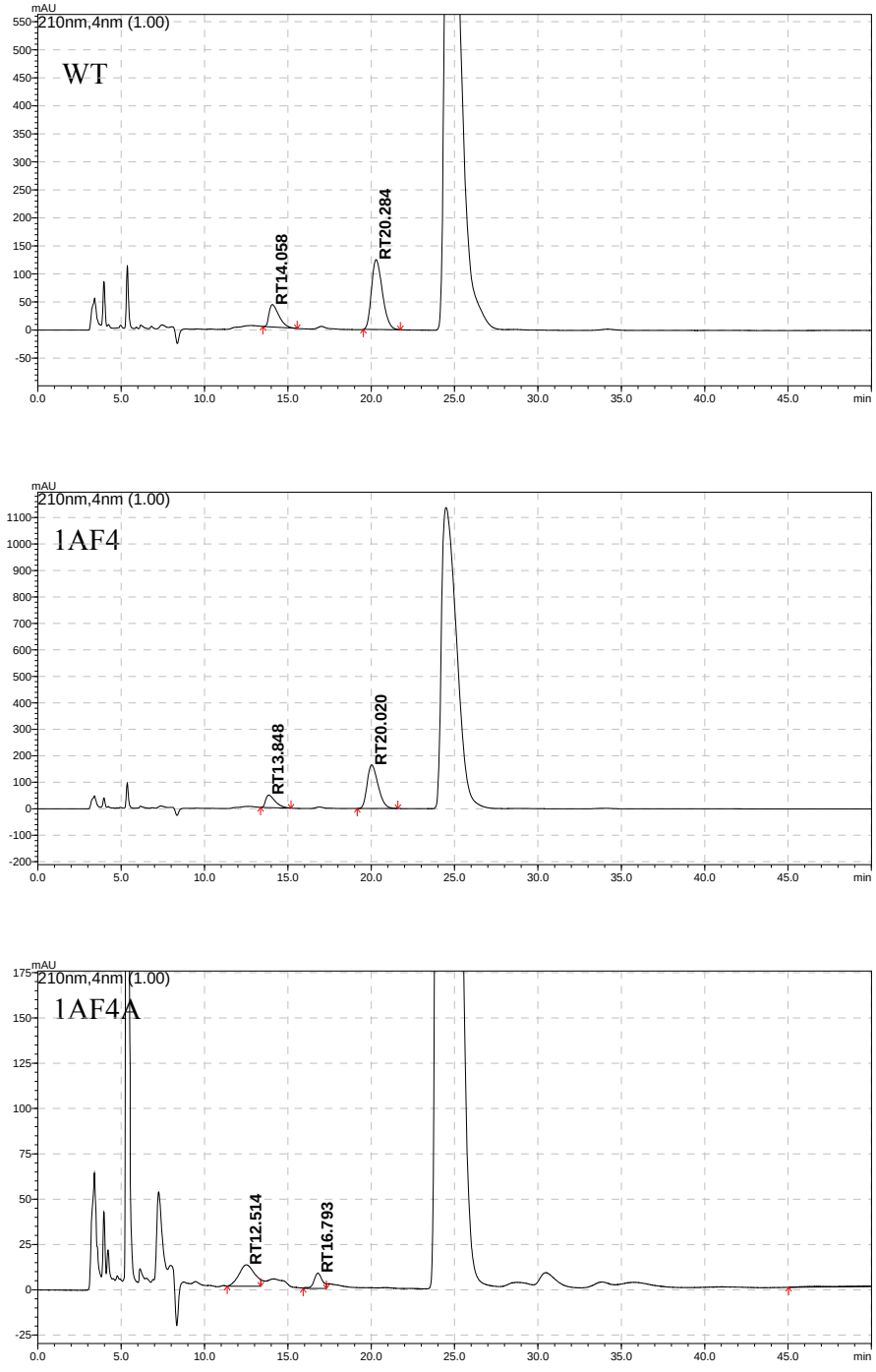


Figure 5.7. Chiral HPLC spectra for the biohydroxylation with substrate **11**.

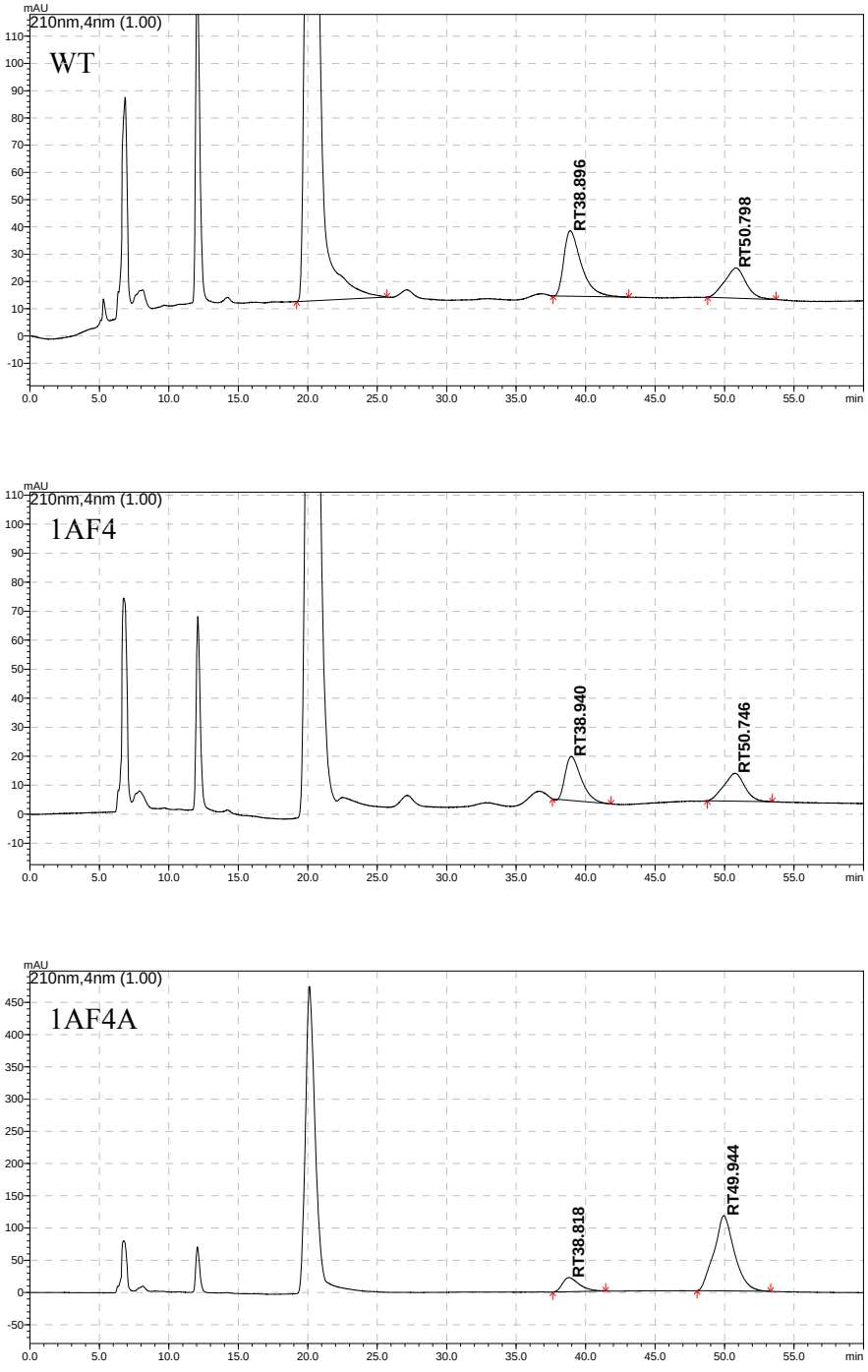


Figure 5.8. Chiral HPLC spectra for the biohydroxylation with substrate **13.**

5.3 Discussion

Six enzymes i.e. WT P450pyr, its mutants 11BB12, 1AF4 and 1AF4A, and its redox partners Fdx and FdR were successfully cloned into pET28a to introduce an N-terminal His_6-tag upstream of the enzyme. The enzymes were expressed in *E. coli* BL21(DE3) and purified using Ni-NTA gravity flow column. For the first time, the P450pyr and its mutants were reconstituted successfully together with their electron transport proteins, Fdx and FdR, and the steady state kinetic parameters were determined. Although the mutants 11BB12, 1AF4 and 1AF4A have higher enantioselectivities, the mutation(s) in the active site region have lowered the enzyme's specific activity and increased its K_M. This is not totally unexpected, and this problem can perhaps be resolved easily with another round of directed evolution (e.g. error-prone PCR) focusing on other areas other than the active pocket to evolve an enzyme which has a higher activity and stability.

When 1AF4 and 1AF4A were used to biohydroxylate different substrates, we found that there was hardly any change in enantioselectivity for substrates **3**, **7** and **9** and only a slight change for substrate **5**. The biohydroxylation of substrate **11** and **13** by 1AF4A showed an inversion of the enantioselectivity. The resistance to changes in enantioselectivity asserts that the enantioselectivity of each substrate is governed by different residues in the active pocket of the P450pyr. Since our P450pyr was engineered using substrate **1**, the selected mutants would result in a significant change of enantioselectivity in substrate **1** but not necessarily other substrates. Different substrates would interact differently with each residue within the active site due to the changes in substrate size as well as the hydrophobic and hydrophilic effect of different directing groups in incoming substrates.

5.4 Conclusion and Outlook

We have successfully reconstituted the P450pyr and its mutants with its electron transport proteins Fdx and FdR *in vitro*. The P450pyr mutants showed poor k_{cat} and K_M which also corresponds to its low activity in whole-cell assays reported in chapter 3. Also, biohydroxylation of the P450pyr mutants with different substrates revealed interesting results, especially for substrates **11** and **13** which showed a reversal in enantioselectivity when biohydroxylated with 1AF4A.

In order to elucidate the relationship and interaction between residues in the P450pyr active pocket and different substrates, it would necessary to crystallize the P450pyr together with its bound substrate. The crystal structure of P450pyr would give fascinating new insights into how the residues within the active pocket could affect the biohydroxylation enantioselectivity of the P450pyr. Future work will focus on performing more detailed characterizations of the P450pyr and its mutants and investigating methods to help increase the activity and enantioselectivity of the P450pyr.

5.5 Materials and Methods

Materials

Escherichia coli BL21(DE3) and the expression plasmids were purchased from Novagen (Madison, WI). PCR grade dNTPs were obtained from Roche Applied Sciences (Indianapolis, IN). Restriction enzymes, Phusion High-Fidelity DNA polymerase, T4 DNA ligase and their corresponding buffers were purchased from

New England Biolabs (NEB) (Beverly, MA). D-glucose was purchased from ThermoFisher (Pittsburgh, PA). Ampicillin, kanamycin, isopropyl *β*-D-thiogalactopyranoside (IPTG), nitro blue tetrazolium (NBT), phenazine methosulfate (PMS), NAD^+, and $NADP^+$ were purchased from Sigma (St. Louis, MO). The Ni-NTA agarose, QIAprep spin plasmid mini-prep kit, QIAEX II gel purification kit, and QIAquick PCR purification kit were purchased from Qiagen (Valencia, CA). Various oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA). SDS-PAGE gels, buffers and protein size markers were purchased from Bio-Rad (Hercules, CA). All other required chemical reagents and solvents were purchased from Acros Organics (Morris Plains, NJ) and Sigma-Aldrich (St. Louis, MO). *N*-benzyloxycarbonyl-pyrrolidine **3** was synthesized according to known procedures.[204]

Cloning of WT P450pyr, 11BB12, 1AF4, 1AF4A, Fdx and FdR

Each gene (WT P450pyr, 11BB12, 1AF4, 1AF4A, Fdx and FdR) was individually amplified using the template and primers listed in Table 5.4. PCRs were carried out in an MJ Research PTC-200 thermal cycler using the following program:

PCR program:

1 = 98 °C for 0:30
2 = 98 °C for 0:10
3 = 55-65 °C for 0:45
4 = 72 °C for 1:00
5 = Goto 2, 24 times
6 = 72 °C for 10:00
7 = 4 °C for ever
8 = End

The PCR products were digested with *Nde*I and *XhoI* or *Hind*III and cloned into their respective sites in pET28a vector in order to express proteins as His_6-tag fusions. The resulting plasmids were transformed *E. coli* DH5α, checked with restriction digestion, sequenced and transformed into *E. coli* BL21(DE3).

Table 5.4. Primers and templates used to amplify 7 different genes

Gene to be amplified	**Primers (Restriction sites underlined)**	**Template**
WT P450pyr	P450 NdeI Fwd: TTA ACT ACT CAT ATG GAA C AT ACA GGA CAA AGC GCG N-P450 XhoI Rev: TTA ACT CTC GAG CTA CGC GTG GAC GCG AAC Primers used in overlap extension PCR: P450 982 fwd: CGCTT GCT CAC ATG CGC CGCA P450 982 rev: TGCG GCG CAT GTG AGC AAGCG	pCom8-PA7F200R1500[131] Note: The NdeI site (CAT ATG) in the P450pyr gene was replaced with CAC ATG by site directed mutagenesis using overlap extension PCR.
11BB12	P450 NdeI Fwd: TTA ACT ACT CAT ATG GAA C AT ACA GGA CAA AGC GCG N-P450 XhoI Rev: TTA ACT CTC GAG CTA CGC GTG GAC GCG AAC	pRSFDuet 11BB12 (from directed evolution step)
1AF4	P450 NdeI Fwd: TTA ACT ACT CAT ATG GAA C AT ACA GGA CAA AGC GCG N-P450 XhoI Rev: TTA ACT CTC GAG CTA CGC GTG GAC GCG AAC	pRSFDuet 1AF4 (from directed evolution step)
1AF4A	P450 NdeI Fwd: TTA ACT ACT CAT ATG GAA C AT ACA GGA CAA AGC GCG N-P450 XhoI Rev: TTA ACT CTC GAG CTA CGC GTG GAC GCG AAC	pRSFDuet 1AF4A (from directed evolution step)
Fdx	Fdx NdeI Fwd: TTA ACT ACT CAT ATG CCA ACA GTG ACC TAT GTT GAA ATA AAT GG Fdx XhoI Rev: TTA ACT ACT	pCom8-PA7F200R1500

Table 5.4 (cont.)

	CTC GAG TCA ATG CTG CGC GAG AGG AAG	
FdR	Red NdeI Fwd: TTA ACT ACT CAT ATG ATC CAC ACC GGC GTG AC Red HindIII Rev: TTA ACT ACT AAG CTT TTA GAG GGA GGT TGG GGA CG	pCom8-PA7F200R1500

Overexpression, and the Purification of WT P450pyr, 11BB12, 1AF4, 1AF4A, Fdx and FdR

E. coli BL21(DE3) cells overexpressing the gene of interest were grown in TB medium supplemented with kanamycin (50 μg mL^{-1}). Cells were grown at 37 °C to OD_{600} of ~0.6 and then induced at 25 °C by the addition of IPTG to a final concentration of 0.25 mM. An additional 0.5 mM of δ-aminolevulinic acid was added to the culture that was expressing P450pyr or its mutants. The culture was incubated at 25 °C for an additional 16 to 18 hours, after which the cells were harvested by centrifugation (10 minutes at 6,000 rpm) and resuspended in cold 10 mM imidazole buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0). Lysozyme (1 mg/mL) was added to the resuspended cells and the sample was freeze-thawed. The cells were further lysed by two passes through the French press. The cell debris was removed by centrifugation (20 minutes at 19,000 rpm, 4 ºC) and the supernatant containing the soluble protein was syringe filtered (Millipore Millex-GP Filter, 0.2 μM).

The N-His_6-tagged-proteins were purified using gravity flow columns containing 3 mL of Qiagen Ni-NTA agarose. The resin was equilibrated with 10 mM imidazole buffer according to the manufacturer's instructions. The clarified supernatant was loaded onto the resin and contaminating *E. coli* proteins were washed with 60 mL of 10 mM imidazole buffer, 60 mL of 20 mM imidazole buffer, 30 mL of 30 mM imidazole buffer and 3 mL of 50 mM imidazole buffer. The targeted proteins were eluted with 200 mM imidazole buffer and the protein fractions were analyzed using the Bradford reagent (Bio-Rad). Fractions containing the protein were pooled together and concentrated a using Millipore Amicon Ultra-4 centrifugal filter device (MWCD 10 kDa) at 4,000 rpm at 4 °C, washed three times with 50 mM potassium phosphate buffer (pH 7.5), and concentrated again. Concentrated protein was stored at -80 °C in 15% glycerol. Protein concentration was determined using the calculated molar extinction coefficient at 280 nm (SDSC Biology Workbench - http://workbench.sdsc.edu/) or by using Bradford protein assay.

CO Difference Spectrum of P450pyr

10 µL of purified P450pyr was diluted 80 fold to a total volume of 800 µL by adding 10 mM potassium phosphate buffer (pH 7.5) in a quartz cuvette. The protein was first reduced by adding 15 mg of sodium hydrosulfite (Sigma-Aldrich) and saturated with CO by bubbling the gas for about 20 sec. The UV absorbance at 450 nm of the protein was measured before and after the CO bubbling. A molar extinction coefficient for P450 of 90 cm^{-1} was used to calculate the P450pyr concentration.

Enzyme Kinetics (*In vitro* Assay)

The enzyme activities of P450pyr, 11BB12, 1AF4 and 1AF4A were routinely assayed by HPLC. The assay mixture (0.5 mL final volume) contained 0.5 μM P450, 5 μM Fdx, 10 μM FdR, substrate **1**, and 1 mM NADH in 50 mM Tris HCl buffer at pH 7.5. Reactions were initiated by the addition of different amounts of substrate **1** to the reaction mixture and incubated at 30 °C. Samples (100 μL) were taken at 10 sec, 1 min, 5 min, and 9 min and were quenched by adding 20 μL of 4 % (v/v) trifluoroacetic acid (TFA). The samples were then analyzed by reverse phase HPLC. The Michaelis-Menten equation was fitted to the data using non-linear least squares regression analysis in Microcal Origin 5.0 (OriginLab Corporation, Northampton, MA) to determine k_{cat} and K_M values.

Reverse Phase HPLC Analysis for *N*-Benzyl pyrrolidine 1 and *N*-Benzyl-3-hydroxypyrrolidines 2

The following conditions were used to analyze samples: Agilent Eclipse Plus C18 (3.5 μm) column (4.6 x 150 mm), UV detection at 210 nm, flow rate of 0.9 mL/min; eluent A, acetonitrile, and eluent B, 10 mM potassium phosphate buffer (pH 7.0). Retention times (min): t_R of **1**, 6.8; t_R of **2**, 10.9, A/B 20:80.

General Procedure for the Biohydroxylation of Different Substrates with Whole-Cells

E. coli cells (BL21(DE3) pRSFDuet P450 pETDuet FR) overexpressing the WT P450 or its mutants (1AF4 and 1AF4A), as well as the Fdx and FdR were grown in 35 mL TB medium supplemented with 50 μg/mL kanamycin and 100 μg/mL

ampicillin. Cells were grown at 37 °C to OD_{600} of ~0.6 and then induced at 30 °C by the addition of IPTG and δ-ALA to a final concentration of 0.5 mM. The culture was incubated at 30 °C for an additional 4 hours, after which the cells were harvested by centrifugation (10 minutes at 4,000 rpm) and resuspended in potassium phosphate buffer (50 mM, pH 7.5) containing 2 % (w/v) glucose to obtain a cell density of 8 g cell dry weight (cdw)/L. 1 mL of cells suspension was aliquoted into different 2 mL microcentrifuge tubes and 4 μL of each substrate (stock solution of 0.5 M) was added into each tube separately. The tubes were shaken at 30 ºC and 1200 rpm for 4 h with a BIOER (Hangzhou, China) mixing block MB-102. The tubes were then centrifuged and the supernatant was extracted for further work-up.

Work-up Procedure for Biohydroxylation of Substrates 1, 3, 5, 7, 9, 11 and 13

A quick and simple work-up procedure was used in this work. The cells were removed from the bioconversion mixture by centrifugation, and the supernatant was adjusted to pH 11-12 by addition of 20 μL of 5 N KOH, followed by extraction with ethyl acetate. The organic phase was separated and the solvent was removed by quick evaporation using a Thermo Scientific SpeedVac concentrator. The product was then resuspended in 100 μL of isopropanol and analyzed with chiral HPLC.

Determination of *ee* of Bioproducts 2, 4, 6, 8, 10, 12 and 14

The following conditions were used to analyze samples: Chiral column (250 m × 4.6 mm). UV detection at 210 and 254 nm, flow rate at 1.0 mL/min (unless stated otherwise); eluent A, *n*-hexane; eluent B, isopropanol. Retention times (min): t_R of (*R*)-**2**, 26.1; t_R of (*S*)-**2**, 43.5, column, chiralcel OB-H (Daicel), A/B 98/2; t_R of (*S*)-**4**, 13.5; t_R of (*R*)-**4**, 15.4, column, chiralpak AS, A/B 9/1; t_R of (*S*)-**6**, 34.8; t_R of (*R*)-**6**,

39, column, chiralpak AS, A/B 99/1; t_R of (*S*)-**8**, 20.3; t_R of (*R*)-**8**, 30.5, column, chiralpak AS, A/B 4/1; t_R of (*R*)-**10**, 17.9; t_R of (*S*)-**10**, 22.6, column, chiralcel OB-H, A/B 7/1, flow 0.5 mL/min; t_R of (*S*)-**12**, 8.7; t_R of (*R*)-**12**, 10.4, column, chiralpak AS, A/B 75/25; t_R of (*S*)-**14**, 28.3; t_R of (*R*)-**14**, 30.8, column, chiralcel OB-H, A/B 95/5, flow 0.5 mL/min.

Chapter 6 : Conclusion and Recommendations

6.1 Conclusion

The recent discovery of a novel P450pyr enzyme from *Sphingomonas* sp. HXN-200, a class I P450 monooxygenase, has led to the discovery of new synthetic applications of biocatalytic hydroxylations of non-activated carbon atoms, as this enzyme is able to catalyze the hydroxylation of a broad range of substrates, such as n-alkanes, cyclic alkanes, *N*-heterocycles, spiro-oxazolidines, and benzoxazole derivatives, with reasonably high activity, regioselectivity and enantioselectivity.[127-129] In this thesis, we seek to investigate and engineer this novel P450pyr at a molecular level in order to further develop our understanding of its biological functions and its potential as an environmental friendly biocatalyst in the chemist's tool box.

There has been many 'firsts' in this thesis. For the first time, a highly active *Escherichia coli* recombinant expressing the P450pyr and its electron transport proteins ferredoxin (Fdx) from *Sphingomonas* sp. HXN200 and the ferredoxin reductase (FdR) from *Mycobacterium* sp. HXN1500 was constructed. We have also developed the first high-throughput colorimetric *ee* assay that was used in the directed evolution of a P450 for asymmetric biohydroxylations. This method was found to be very sensitive, independent of concentration, and takes only a few seconds to measure each 96-well microtiter plate. In addition, we have developed a high-throughput mass spectrometry (MS)-based screening method which has high accuracy in *ee* determination and can be used for different hydroxylation reactions, independent of the biocatalyst. Other advantages of this MS-based method are its high sensitivity,

speed, and simplicity as it allows the direct analysis of samples taken from the supernatant of biohydroxylation reactions.

We have also reported the first example where directed evolution, coupled with a high-throughput enantioselectivity screening system, has been applied to improve the biohydroxylation enantioselectivity of a class I P450 monooxygenase using a prochiral substrate. We have succeeded in inverting the enantioselectivity of the P450pyr towards the biohydroxylation of prochiral substrate *N*-benzyl-3-hydroxypyrrolidine by iterative targeted site-saturation mutagenesis coupled with a high-throughput two-enzyme-based colorimetric assay. An interesting observation to note was that one mutation at position 100 was all it took to completely invert the enantioselectivity of the P450pyr. The best mutant 1AF4A has an *ee* of 83% (*R*) compared to the wild type's *ee* of 43% (*S*).

In order to increase the efficiency and productivity of the P450pyr-catalyzed hydroxylation reaction, two cofactor regeneration systems utilizing either the glucose dehydrogenase[118] (GDH) from Bacillus substilis BGSC 1A1 or the engineered phosphite dehydrogenase 12x[33,123] (PTDH 12x) from *Pseudomonas stutzeri* were incorporated together with the P450pyr system *in vivo*. As a result, we have developed several strains of *E. coli* which have shown up to 2.5x improvement in specific activity compared to the parent strain which did not have the cofactor-regeneration system. The phosphite dehydrogenase-based cofactor regeneration system was also found to have better pH-buffering since the side product of the phosphite dehydrogenase system is phosphate which can act as a buffer while the glucose dehydrogenase produces an acidic gluconolactone.

6.2 Recommendations/ Future Work

Although we have succeeded in improving the enantioselectivity of the P450pyr, the activity of the enzyme was negatively affected. This is not totally unexpected as the introduction of mutations to its heme active site by iterative targeted site-saturation mutagenesis has probably introduced a destabilizing effect in the protein. Hence, it would be necessary to take a step back to increase the P450pyr stability and activity by random mutagenesis before attempting to proceed with further rounds of mutagenesis to improve the enzyme's enantioselectivity. With increased stability, the enzyme would have more room to evolve as it can tolerate highly destabilizing mutations.[174].

Using our directed evolution approach, we have also discovered the (*S*)-enantioselective mutant 11BB12 that has an *ee* of 65% (*S*). It would be interesting if we could utilize a different directed evolution approach, e.g. CASTing[205] or gene shuffling, coupled with the high-throughput mass spectrometry-based screening method to evolve an (*S*)-enantioselective mutant with a much higher enantioselectivity (i.e. >90%)

In order to further characterize our P450pyr mutants, it would be necessary to crystallize the P450pyr mutants together with their bound substrate. The crystal structure of wild type P450pyr and its mutants would give us a deeper understanding of the structure-function relationship of the residues in the P450pyr active pocket and enable us to predict how certain residues within the active pocket could affect the biohydroxylation enantioselectivity of the P450pyr.

The *in vivo* cofactor regeneration system which we have developed can also be applied to the production of different substrates or used in other microbial biohydroxylation reactions. We could also explore different promoters, plasmid and gene combinations, as well as different reaction conditions to further improve the productivity of the biohydroxylation system. In a different approach, we could utilize the enzymes co-immobilization on nanoparticles method[206] to further improve the productivity of the P450pyr system. The advantages of this nanoparticle immobilization system include recyclability of nanoparticles, high dispersity, larger enzyme loading amounts and lower mass transfer resistance, thus leading to higher turnover rates.

References

1. Soetaert, W. & Vandamme, E. The impact of industrial biotechnology. *Biotechnol J* **1**, 756-769 (2006).

2. Lundy, D., Nesbitt, E. & Polly, L. Development & adoption of industrial biotechnology by the US chemical & biofuel industries. *Ind Biotechnol* **4**, 262-287 (2008).

3. Lorenz, P. & Zinke, H. White biotechnology: Differences in US and EU approaches? *Trends Biotechnol* **23**, 570-574 (2005).

4. Sijbesma, F. Industrial (white) biotechnology: An effective route to increase EU innovation and sustainable growth. in *DSM Report* (2004).

5. Nesbitt, E.R. Industrial biotechnology in China amidst changing market conditions. *J Int Commerce and Economics*, http://www.usitc.gov/journal/documents/biotechnology_china.pdf (2009).

6. Carey, J.S., Laffan, D., Thomson, C. & Williams, M.T. Analysis of the reactions used for the preparation of drug candidate molecules. *Org Biomol Chem* **4**, 2337-2347 (2006).

7. Thoden, J.B. et al. Evolution of enzymatic activity in the enolase superfamily: Structural studies of the promiscuous o-succinylbenzoate synthase from Amycolatopsis. *Biochemistry* **43**, 5716-5727 (2004).

8. Straathof, A.J.J., Panke, S. & Schmid, A. The production of fine chemicals by biotransformations. *Curr Opin Biotechnol* **13**, 548-556 (2002).

9. Labinger, J.A. & Bercaw, J.E. Understanding and exploiting C-H bond activation. *Nature* **417**, 507-514 (2002).

10. Li, Z. & Chang, D.L. Recent advances in regio- and stereoselective biohydroxylation of non-activated carbon atoms. *Curr Org Chem* **8**, 1647-1658 (2004).

11. Bernhardt, R. Cytochromes P450 as versatile biocatalyst. *J Biotechnol* **124**, 128-145 (2006).

12. Sono, M., Roach, M.P., Coulter, E.D. & Dawson, J.H. Heme-containing oxygenases. *Chem Rev* **96**, 2841-2887 (1996).

13. Ro, D.K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-3 (2006).

14. Lieberman, S. & Lin, Y.Y. Reflections on sterol sidechain cleavage process catalyzed by cytochrome $P450_{scc}$. *J Steroid Biochem Mol Biol* **78**, 1-14 (2001).

15. Rabe, K., Gandubert, V., Spengler, M., Erkelenz, M. & Niemeyer, C. Engineering and assaying of cytochrome P450 biocatalysts. *Anal Bioanal Chem* **392**, 1059-1073 (2008).

16. Gillam, E.M. Engineering cytochrome P450 enzymes. *Chem Res Toxicol* **21**, 220-31 (2008).

17. Elliott, S.J. et al. Regio- and stereoselectivity of particulate methane monooxygenase from *Methylococcus capsulatus* (Bath). *J Am Chem Soc* **119**, 9949-9955 (1997).

18. Hakemian, A.S. & Rosenzweig, A.C. The biochemistry of methane oxidation *Annu Rev Biochem* **76**, 223-241 (2007).

19. Liu, K.E., Johnson, C.C., Newcomb, M. & Lippard, S.J. Radical clock substrate probes and kinetic isotope effect studies of the hydroxylation of hydrocarbons by methane monooxygenase. *J Am Chem Soc* **115**, 939-947 (1993).

20. van Beilen, J.B. et al. Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: Evolution and regulation of the *alk*-genes. *Microbiology* **147**, 1621-30 (2001).

21. van Beilen, J.B. & Funhoff, E.G. Expanding the alkane oxygenase toolbox: New enzymes and applications. *Curr Opin Biotechnol* **16**, 308-14 (2005).

22. van Beilen, J.B. et al. Identification of an amino acid position that determines the substrate range of integral membrane alkane hydroxylases. *J Bacteriol* **187**, 85-91 (2005).

23. Luetz, S., Giver, L. & Lalonde, J. Engineered enzymes for chemical production. *Biotechnol Bioeng* **101**, 647-653 (2008).

24. Martí, S. et al. Computational design of biological catalysts. *Chem Soc Rev* **37**, 2634-2643 (2008).

25. Ward, T.R. Artificial enzymes made to order: Combination of computational design and directed evolution. *Angew Chem Int Ed* **47**, 7802-7803 (2008).

26. Sterner, R., Merkl, R. & Raushel, F.M. Computational design of enzymes. *Chem Biol* **15**, 421-423 (2008).

27. Graham-Lorence, S. et al. An active site substitution, F87V, converts cytochrome P450 BM-3 into a regio- and stereoselective (14S,15R)-arachidonic acid epoxygenase. *J Biol Chem* **272**, 1127-35 (1997).

28. Sulistyaningdyah, W.T. et al. Hydroxylation activity of P450 BM-3 mutant F87V towards aromatic compounds and its application to the synthesis of hydroquinone derivatives from phenolic compounds. *Appl Microbiol Biotechnol* **67**, 556-62 (2005).

29. Kumar, S., Scott, E.E., Liu, H. & Halpert, J.R. A rational approach to re-engineer cytochrome P450 2B1 regioselectivity based on the crystal structure of cytochrome P450 2C5. *J Biol Chem* **278**, 17178-84 (2003).

30. Sheldon, R.A. & van Rantwijk, F. Biocatalysis for sustainable organic synthesis. *Australian J Chem* **57**, 281-289 (2004).

31. Eijsink, V.G.H., Gaseidnes, S., Borchert, T.V. & van den Burg, B. Directed evolution of enzyme stability. *Biomol Eng* **22**, 21-30 (2005).

32. Hibbert, E.G. et al. Directed evolution of biocatalytic processes. *Biomol Eng* **22**, 11-19 (2005).

33. Johannes, T.W., Woodyer, R.D. & Zhao, H. Directed evolution of a thermostable phosphite dehydrogenase for NAD(P)H regeneration. *Appl Environ Microbiol* **71**, 5728-5734 (2005).

34. Stemmer, W.P.C. DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. *Proc Nat Acad Sci USA* **91**, 747-751 (1994).

35. Zhao, H., Giver, L., Shao, Z., Affholter, J.A. & Arnold, F.H. Molecular evolution by staggered extention process (StEP) in vitro recombination. *Nat Biotechnol* **16**, 258-261 (1998).

36. Kubo, T., Peters, M.W., Meinhold, P. & Arnold, F.H. Enantioselective epoxidation of terminal alkenes to (*R*)- and (*S*)-epoxides by engineered cytochromes P450 BM-3. *Chem Eur J* **12**, 1216-1220 (2006).

37. Chockalingam, K., Chen, Z.L., Katzenellenbogen, J.A. & Zhao, H.M. Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc Nat Acad Sci USA* **102**, 5691-5696 (2005).

38. Chockalingam, K. & Zhao, H. Creating new specific ligand-receptor pairs for transgene regulation. *Trends Biotechnol* **23**, 333-5 (2005).

39. Islam, K.M.D. et al. Directed evolution of estrogen receptor proteins with altered ligand-binding specificities. *Protein Eng Des Sel* **22**, 45-52 (2009).

40. Nair, N.U. & Zhao, H. Evolution in reverse: Engineering a D-xylose-specific xylose reductase. *Chembiochem* **9**, 1213-5 (2008).

41. Zha, W., Rubin-Pitel, S. & Zhao, H. Exploiting genetic dversity by directed evolution: Molecular breeding of Type III polyketide synthases improves productivity. *Mol BioSyst* **4**, 246-248 (2008).

42. Rubin-Pitel, S., Cho, C.M.-H., Chen, W. & Zhao, H. Directed evolution tools in bioproduct and bioprocess development. in *Bioprocessing for Value-Added Products from Renewable Resources: New Technologies and Applications* 49-72 (Elsevier Science, New York, 2006).

43. Glieder, A., Farinas, E.T. & Arnold, F.H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nat Biotechnol* **20**, 1135-9 (2002).

44. Fasan, R., Meharenna, Y.T., Snow, C.D., Poulos, T.L. & Arnold, F.H. Evolutionary history of a specialized P450 propane monooxygenase. *J Mol Biol* **383**, 1069-80 (2008).

45. Lentz, O. et al. Altering the regioselectivity of cytochrome P450 CYP102A3 of *Bacillus subtilis* by using a new versatile assay system. *Chembiochem* **7**, 345-50 (2006).

46. Wong, T.S., Arnold, F.H. & Schwaneberg, U. Laboratory evolution of cytochrome P450 BM-3 monooxygenase for organic cosolvents. *Biotechnol Bioeng* **85**, 351-8 (2004).

47. Salazar, O., Cirino, P.C. & Arnold, F.H. Thermostabilization of a cytochrome P450 peroxygenase. *Chembiochem* **4**, 891-3 (2003).

48. Boersma, Y.L. et al. A novel genetic selection system for improved enantioselectivity of *Bacillus subtilis* lipase A. *Chembiochem* **9**, 1110-1115 (2008).

49. Leemhuis, H., Kelly, R.M. & Dijkhuizen, L. Directed evolution of enzymes: Library screening strategies. *IUBMB Life* **61**, 222-228 (2009).

50. McLachlan, M., Sullivan, R.P. & Zhao, H. Directed enzyme evolution and high throughput screening. in *Biocatalysis for the Pharmaceutical Industry-Discovery, Development, and Manufacturing* (eds. Tao, J., Lin, G. & Liese, A.) 45-64 (John Wiley and Sons, Singapore, 2009).

51. Reetz, M.T., Zonta, A., Schimossek, K., Jaeger, K.E. & Liebeton, K. Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution. *Angew Chem Int Ed* **36**, 2830-2832 (1997).

52. Janes, L.E. & Kazlauskas, R.J. Quick E. A fast spectrophotometric method to measure the enantioselectivity of hydrolases. *J Org Chem* **62**, 4560-4561 (1997).

53. Janes, L.E., Löwendahl, A.C. & Kazlauskas, R.J. Quantitative screening of hydrolase libraries using pH indicators: Identifying active and enantioselective hydrolases. *Chem Eur J* **4**, 2324-2331 (1998).

54. Tumambac, G.E. & Wolf, C. Enantioselective analysis of an asymmetric reaction using a chiral fluorosensor. *Org Lett* **7**, 4045-4048 (2005).

55. Hwang, B.-Y. & Kim, B.-G. High-throughput screening method for the identification of active and enantioselective ω-transaminases. *Enz Microb Technol* **34**, 429-436 (2004).

56. Reetz, M.T. & Ruggeberg, C.J. A screening system for enantioselective enzymes based on differential cell growth. *Chem Commun*, 1428-1429 (2002).

57. Reetz, M.T., Becker, M.H., Klein, H.-W. & Stöckigt, D. A method for high-throughput screening of enantioselective catalysts. *Angew Chem Int Ed* **38**, 1758-1761 (1999).

58. Cedrone, F. et al. Directed evolution of the epoxide hydrolase from *Aspergillus niger*. *Biocatal Biotransform* **21**, 357 - 364 (2003).

59. Reetz, M., Eipper, A., Tielmann, P. & Mynott, R. A practical NMR-based high-throughput assay for screening enantioselective catalysts and biocatalysts. *Adv Synth Catal* **344**, 1008-1016 (2002).

60. Tielmann, P., Boese, M., Luft, M. & Reetz, M.T. A practical high-throughput screening system for enantioselectivity by using FTIR spectroscopy. *Chem Eur J* **9**, 3882-3887 (2003).

61. DeSantis, G. et al. Creation of a productive, highly enantioselective nitrilase through Gene Site Saturation Mutagenesis (GSSM). *J Am Chem Soc* **125**, 11476-11477 (2003).

62. Abato, P. & Seto, C.T. EMDee: An enzymatic method for determining enantiomeric excess. *J Am Chem Soc* **123**, 9206-9207 (2001).

63. Taran, F. et al. High-throughput screening of enantioselective catalysts by immunoassay. *Angew Chem Int Ed* **41**, 124-127 (2002).

64. Belder, D., Ludwig, M., Wang, L.W. & Reetz, M.T. Enantioselective catalysis and analysis on a chip. *Angew Chem Int Ed* **45**, 2463-2466 (2006).

65. Wong, C.H. & Whitesides, G.M. *Enzymes in synthetic organic chemistry*, (Elsevier Science Ltd, Oxford, UK, 1994).

66. Liese, A. & Filho, M.V. Production of fine chemicals using biocatalysis. *Curr Opin Biotechnol* **10**, 595-603. (1999).

67. Faber, K. *Biotransformations in Organic Chemistry*, 454 (Springer Verlag, Berlin, Germany, 2004).

68. Hummel, W. & Kula, M.R. Dehydrogenases for the synthesis of chiral compounds. *Eur J Biochem* **184**, 1-13 (1989).

69. Hummel, W. New alcohol dehydrogenases for the synthesis of chiral compounds. *Adv Biochem Eng Biotechnol* **58**, 145-84 (1997).

70. Kirk, O., Borchert, T.V. & Fuglsang, C.C. Industrial enzyme applications. *Curr Opin Biotechnol* **13**, 345-351 (2002).

71. Koeller, K.M. & Wong, C.H. Enzymes for chemical synthesis. *Nature* **409**, 232-40. (2001).

72. Li, Z. et al. Oxidative biotransformations using oxygenases. *Curr Opin Chem Biol* **6**, 136-44 (2002).

73. Stewart, J.D. Dehydrogenases and transaminases in asymmetric synthesis. *Curr Opin Chem Biol* **5**, 120-9 (2001).

74. Jossek, R. & Steinbuchel, A. In vitro synthesis of poly(3-hydroxybutyric acid) by using an enzymatic coenzyme A recycling system. *FEMS Microbiol Lett* **168**, 319-324 (1998).

75. Satoh, Y., Tajima, K., Tannai, H. & Munekata, M. Enzyme-catalyzed poly(3-hydroxybutyrate) synthesis from acetate with CoA recycling and NADPH regeneration in vitro. *J Biosci Bioeng* **95**, 335-341 (2003).

76. Patel, S.S., Conlon, H.D. & Walt, D.R. Enzyme catalyzed synthesis of L-acetylcarnitine and citric acid using acetyl coenzyme-A recycling. *J Org Chem* **51**, 2842-2844 (1986).

77. Ouyang, T.M., Walt, D.R. & Patel, S.S. Enzyme catalyzed synthesis of citric acid using acetyl coenzyme A recycling in a 2-phase system. *Bioorg Chem* **18**, 131-135 (1990).

78. Chenault, H., Simon, E. & Whitesides, G. Cofactor regeneration for enzyme-catalysed synthesis. *Biotechnol Genet Eng Rev* **6**, 221-270 (1988).

79. Adlercreutz, P. Cofactor regeneration in biocatalysis in organic media *Biocatal Biotransform* **14**, 1-30 (1996).

80. Chenault, H. & Whitesides, G. Regeneration of nicotinamide cofactors for use in organic synthesis. *Appl Biochem Biotechnol* **14**, 147-197 (1987).

81. Eckstein, M., Daumann, T. & Kragl, U. Recent developments in NAD(P)H regeneration for enzymatic reductions in one- and two-phase systems *Biocatal Biotransform* **22**, 89-96 (2004).

82. Goldberg, K., Schroer, K., Lutz, S. & Liese, A. Biocatalytic ketone reduction--a powerful tool for the production of chiral alcohols--part I: processes with isolated enzymes. *Appl Microbiol Biotechnol* **76**, 237-48 (2007).

83. Hollmann, F. & Schmid, A. Electrochemical regeneration of oxidoreductases for cell-free biocatalytic redox reactions. *Biocatal Biotransform* **22**, 63 - 88 (2004).

84. Hummel, W. Large-scale applications of NAD(P)-dependent oxidoreductases: Recent developments. *Trends Biotechnol* **17**, 487-492 (1999).

85. Kroutil, W., Mang, H., Edegger, K. & Faber, K. Recent advances in the biocatalytic reduction of ketones and oxidation of sec-alcohols. *Curr Opin Chem Biol* **8**, 120-6 (2004).

86. van der Donk, W.A. & Zhao, H. Recent developments in pyridine nucleotide regeneration. *Curr Opin Biotechnol* **14**, 421-6 (2003).

87. Zhao, H. & van der Donk, W.A. Regeneration of cofactors for use in biocatalysis. *Curr Opin Biotechnol* **14**, 583-589 (2003).

88. Ouyang, T. & Walt, D.R. A new chemical method for synthesizing and recycling acyl coenzyme A thioesters. *J Org Chem* **56**, 3752-3755 (1991).

89. Boonstra, B., Rathbone, D.A., French, C.E., Walker, E.H. & Bruce, N.C. Cofactor regeneration by a soluble pyridine nucleotide transhydrogenase for biological production of hydromorphone. *Appl Environ Microbiol* **66**, 5161-6. (2000).

90. Berrios-Rivera, S.J., Bennett, G.N. & San, K.Y. Metabolic engineering of *Escherichia coli*: increase of NADH availability by overexpressing an NAD(+)-dependent formate dehydrogenase. *Metab Eng* **4**, 217-29. (2002).

91. Eguchi, T. et al. NADPH regeneration by glucose dehydrogenase from Gluconobacter scleroides for l-leucovorin synthesis. *Biosci Biotechnol Biochem* **56**, 701-3 (1992).

92. Crans, D.C. & Whitesides, G.M. A convenient synthesis of disodium acetyl phosphate for use in *in situ* ATP cofactor regeneration. *J Org Chem* **48**, 3130-3132 (1983).

93. Hirschbein, B.L., Mazenod, F.P. & Whitesides, G.M. Synthesis of phosphoenolpyruvate and its use in adenosine-triphosphate cofactor regeneration. *J Org Chem* **47**, 3765-3766 (1982).

94. Hoffman, R.C. et al. Immobilized polyphosphate kinase - preparation, properties, and potential for use in adenosine 5'-triphosphate regeneration. *Biotechnol Appl Biochem* **10**, 107-117 (1988).

95. Kim, D.M. & Swartz, J.R. Regeneration of adenosine triphosphate from glycolytic intermediates for cell-free protein synthesis. *Biotechnol Bioeng* **74**, 309-16 (2001).

96. Kondo, H., Tomioka, I., Nakajima, H. & Imahori, K. Construction of a system for the regeneration fo ATP, which supplies energy to bioreactor. *J Appl Biohem* **6**, 29-38 (1984).

97. Resnick, S.M. & Zehnder, A.J. In vitro ATP regeneration from polyphosphate and AMP by polyphosphate:AMP phosphotransferase and adenylate kinase from *Acinetobacter johnsonii* 210A. *Appl Environ Microbiol* **66**, 2045-51 (2000).

98. Chen, X. et al. Sugar nucleotide regeneration beads (superbeads): A versatile tool for the practical synthesis of oligosaccharides. *J Am Chem Soc* **123**, 2081-2082 (2001).

99. Bulter, T. & Elling, L. Enzymatic synthesis of nucleotide sugars. *Glycoconj J* **16**, 147-59 (1999).

100. Fujio, T. & Maruyama, A. Enzymatic production of pyrimidine nucleotides using *Corynebacterium ammoniagenes* cells and recombinant *Escherichia coli* cells: Enzymatic production of CDP-choline from orotic acid and choline chloride .1. *Biosci Biotechnol Biochem* **61**, 956-959 (1997).

101. Burkart, M.D., Izumi, M. & Wong, C.H. Enzymatic regeneration of 3 '-phosphoadenosine-5 '-phosphosulfate using aryl sulfotransferase for the preparative enzymatic synthesis of sulfated carbohydrates. *Angew Chem Int Ed* **38**, 2747-2750 (1999).

102. Burkart, M.D., Izumi, M., Chapman, E., Lin, C.H. & Wong, C.H. Regeneration of PAPS for the enzymatic synthesis of sulfated oligosaccharides. *J Org Chem* **65**, 5565-5574 (2000).

103. Chen, X. et al. Transferring a biosynthetic cycle into a productive *Escherichia coli* strain: Large-scale synthesis of galactosides. *J Am Chem Soc* **123**, 8866-8867 (2001).

104. Koizumi, S., Endo, T., Tabata, K. & Ozaki, A. Large-scale production of UDP-galactose and globotriose by coupling metabolically engineered bacteria. *Nat Biotechnol* **16**, 847-50 (1998).

105. Woltinger, J., Karau, A., Leuchtenberger, W. & Drauz, K. Membrane reactors at Degussa. *Adv Biochem Eng Biotechnol* **92**, 289-316 (2005).

106. Kragl, U., Kruse, W., Hummel, W. & Wandrey, C. Enzyme engineering aspects of biocatalysis: Cofactor regeneration as example. *Biotechnol Bioeng* **52**, 309-319 (1996).

107. McCoy, M. Making Drugs with Little Bugs. *Chem Eng News* **79**, 37-43 (2001).

108. Slusarczyk, H., Felber, S., Kula, M.R. & Pohl, M. Stabilization of NAD-dependent formate dehydrogenase from *Candida boidinii* by site-directed mutagenesis of cysteine residues. *Eur J Biochem* **267**, 1280-1289 (2000).

109. Hollmann, F., Schmid, A. & Steckhan, E. The first synthetic application of a monooxygenase employing indirect electrochemical NADH regeneration. *Angew Chem Int Ed* **40**, 169-171 (2001).

110. Hollmann, F., Witholt, B. & Schmid, A. [Cp*Rh(bpy)(H2O)](2+): a versatile tool for efficient and non-enzymatic regeneration of nicotinamide and flavin coenzymes. *J. Mol. Cat. B: Enzymatic* **19**, 167-176 (2002).

111. Antiochia, R., Lavagnini, I. & Magno, F. Electrocatalytic oxidation of NADH at single-wall carbon-nanotube-paste electrodes: kinetic considerations for use of a redox mediator in solution and dissolved in the paste. *Anal Bioanal Chem* **381**, 1355-61 (2005).

112. Jiang, Z.Y., Lu, C.Q. & Wu, H. Photoregeneration of NADH using carbon-containing TiO2. *Ind Eng Chem Research* **44**, 4165-4170 (2005).

113. Lo, H.C. et al. Bioorganometallic chemistry. 13. Regioselective reduction of NAD(+) models, 1-benzylnicotinamde triflate and beta-nicotinamide ribose-5 '-methyl phosphate, with in situ generated [CP*Rh(Bpy)H](+): Structure-activity relationships, kinetics, and mechanistic aspects in the formation of the 1,4-NADH derivatives. *Inorganic Chemistry* **40**, 6705-6716 (2001).

114. Matsuo, T. & Mayer, J.M. Oxidations of NADH analogues by cis-$[RuIV(bpy)_2(py)(O)]^{2+}$ occur by hydrogen-atom transfer rather than by hydride transfer. *Inorg Chem* **44**, 2150-8 (2005).

115. Westerhausen, D., Herrmann, S., Hummel, W. & Steckhan, E. Formate-driven, nonenzymic NAD(P)H regeneration in the alcohol dehydrogenase-catalyzed stereoselective reduction of 4-phenyl-2-butanone. *Angew Chem Int Ed* **31**, 1529-31 (1992).

116. Wagenknecht, P.S., Penney, J.M. & Hembre, R.T. Transition-Metal-Catalyzed Regeneration of Nicotinamide Coenzymes with Hydrogen. *Organometallics* **22**, 1180-1182 (2003).

117. Devaux-Basséguy, R., Bergel, A. & Comtat, M. Potential Applications of NAD(P)-Dependent Oxidoreductases in Synthesis - a Survey. *Enz Microb Technol* **20**, 248-258 (1997).

118. Zhang, W., O'Connor, K., Wang, D.I. & Li, Z. Bioreduction with efficient recycling of NADPH by coupled permeabilized microorganisms. *Appl Environ Microbiol* **75**, 687-94 (2009).

119. Betancor, L., Berne, C., Luckarift, H.R. & Spain, J.C. Coimmobilization of a redox enzyme and a cofactor regeneration system. *Chem Commun*, 3640-3642 (2006).

120. Costas, A.M., White, A.K. & Metcalf, W.W. Purification and characterization of a novel phosphorus-oxidizing enzyme from *Pseudomonas stutzeri* WM88. *J Biol Chem* **276**, 17429-36. (2001).

121. Vrtis, J.M., White, A.K., Metcalf, W.W. & van der Donk, W.A. Phosphite dehydrogenase: a versatile cofactor-regeneration enzyme. *Angew Chem Int Ed* **41**, 3257-9 (2002).

122. Woodyer, R., van der Donk, W.A. & Zhao, H. Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. *Biochemistry* **42**, 11604-11614 (2003).

123. Johannes, T.W., Woodyer, R.D. & Zhao, H. Efficient regeneration of NADPH using an engineered phosphite dehydrogenase. *Biotechnol Bioeng* **96**, 18-26 (2007).

124. Clerval, R. et al. The come-back of high-throughput screening of wild-type microbial strains through the use of miniaturised growth systems and LC-MS. *Bioworld* **6**, 24-26 (2000).

125. Duetz, W.A. et al. Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates. *Appl Environ Microbiol* **66**, 2641-2646 (2000).

126. Li, Z., Feiten, H.-J., van Beilen, J.B., Duetz, W. & Witholt, B. Preparation of optically active *N*-benzyl-3-hydroxypyrrolidine by enzymatic hydroxylation. *Tet Asymm* **10**, 1323-1333 (1999).

127. Chang, D. et al. Practical syntheses of *N*-substituted 3-hydroxyazetidines and 4-hydroxypiperidines by hydroxylation with *Sphingomonas* sp. HXN-200. *Org Lett* **4**, 1859-1862 (2002).

128. Chang, D., Feiten, H.-J., Witholt, B. & Li, Z. Regio- and stereoselective hydroxylation of *N*-substituted piperidin-2-ones with *Sphingomonas* sp. HXN-200. *Tet Asymm* **13**, 2141-2147 (2002).

129. Chang, D., Witholt, B. & Li, Z. Preparation of (*S*)-*N*-substituted 4-hydroxy-pyrrolidin-2-ones by regio- and stereoselective hydroxylation with *Sphingomonas* sp. HXN-200. *Org Lett* **2**, 3949-3952 (2000).

130. Li, Z. et al. Preparation of (*R*)- and (*S*)-*N*-protected 3-hydroxypyrrolidines by hydroxylation with *Sphingomonas* sp. HXN-200, a highly active, regio- and stereoselective, and easy to handle biocatalyst. *J Org Chem* **66**, 8424-30 (2001).

131. van Beilen, J.B. et al. Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases. *Appl Environ Microbiol* **72**, 59-65 (2006).

132. Reetz, M.T. Combinatorial and evolution-based methods in the creation of enantioselective catalysts. *Angew Chem Int Ed* **40**, 284-310 (2001).

133. Wahler, D. & Reymond, J.L. High-throughput screening for biocatalysts. *Curr Opin Biotechnol* **12**, 535-44 (2001).

134. Finn, M.G. Emerging methods for the rapid determination of enantiomeric excess. *Chirality* **14**, 534-40 (2002).

135. Reetz, M.T. New methods for the high-throughput screening of enantioselective catalysts and biocatalysts. *Angew Chem Int Ed* **41**, 1335-8 (2002).

136. Schmidt, M. & Bornscheuer, U.T. High-throughput assays for lipases and esterases. *Biomol Eng* **22**, 51-6 (2005).

137. Muller, C.A., Markert, C., Teichert, A.M. & Pfaltz, A. Mass spectrometric screening of chiral catalysts and catalyst mixtures. *Chem Commun*, 1607-18 (2009).

138. Otten, L.G., Hollmann, F. & Arends, I.W. Enzyme engineering for enantioselectivity: from trial-and-error to rational design? *Trends Biotechnol* **28**, 46-54 (2010).

139. Reetz, M.T. et al. A GC-based method for high-throughput screening of enantioselective catalysts. *Catal Today* **67**, 389-396 (2001).

140. Ding, K., Ishii, A. & Mikami, K. Super high throughput screening (SHTS) of chiral ligands and activators: Asymmetric activation of chiral diol-zinc catalysts by chiral nitrogen activators for the enantioselective addition of diethylzinc to aldehydes. *Angew Chem Int Ed* **38**, 497-501 (1999).

141. Mikami, K. et al. Asymmetric activation of chiral alkoxyzinc catalysts by chiral nitrogen activators for dialkylzinc addition to aldehydes: super high-throughput screening of combinatorial libraries of chiral ligands and activators by HPLC-CD/UV and HPLC-OR/RIU systems. *Chemistry* **7**, 730-7 (2001).

142. Manfred , T.R., Klaus , M.K., Alfred, D., Heike, H. & Detlev, B. Super-high-throughput screening of enantioselective catalysts by using capillary array electrophoresis. *Angew Chem Int Ed* **39**, 3891-3893 (2000).

143. Tao, W.A. & Cooks, R.G. Parallel reactions for enantiomeric quantification of peptides by mass spectrometry. *Angew Chem Int Ed* **40**, 757-760 (2001).

144. Richard, A.v.D. & Ben, L.F. Color indicators of molecular chirality based on doped liquid crystals. *Angew Chem Int Ed* **40**, 3198-3200 (2001).

145. Frédéric, T. et al. High-throughput screening of enantioselective catalysts by immunoassay. *Angew Chem Int Ed* **41**, 124-127 (2002).

146. Schoofs, A. & Horeau, A. Nouvelle methode generale de determination de la purete enantiomerique et de la configuration absolue des alcools secondaires chiraux. *Tet Letts* **18**, 3259-3262 (1977).

147. Guo, J., Wu, J., Siuzdak, G. & Finn, M.G. Measurement of enantiomeric excess by kinetic resolution and mass spectrometry. *Angew Chem Int Ed* **38**, 1755-1758 (1999).

148. Yao, S., Meng, J.-C., Siuzdak, G. & Finn, M.G. New catalysts for the asymmetric hydrosilylation of ketones discovered by mass spectrometry screening. *J Org Chem* **68**, 2540-2546 (2003).

149. Korbel, G.A., Lalic, G. & Shair, M.D. Reaction microarrays: A method for rapidly determining the enantiomeric excess of thousands of samples. *J Am Chem Soc* **123**, 361-362 (2000).

150. Millot, N. et al. Rapid determination of enantiomeric excess using infrared thermography. *Org Proc Res Dev* **6**, 463-470 (2002).

151. Li, Z., Butikofer, L. & Witholt, B. High-throughput measurement of the enantiomeric excess of chiral alcohols by using two enzymes. *Angew Chem Int Ed* **43**, 1698-702 (2004).

152. Tang, W.L., Li, Z. & Zhao, H. Inverting the enantioselectivity of P450pyr monooxygenase by directed evolution. *Chem Commun* **46**, 5461-5463 (2010).

153. Chen, Y., Tang, W.L., Mou, J. & Li, Z. High-throughput method for determining the enantioselectivity of enzyme-catalyzed hydroxylations based on mass spectrometry. *Angew Chem Int Ed* **49**, 5278-5283 (2010).

154. Dey, S., Powell, D.R., Hu, C. & Berkowitz, D.B. "Cassette" in situ enzymatic screening identifies complementary chiral scaffolds for hydrolytic kinetic resolution across a range of epoxides. *Angew Chem Int Ed* **46**, 7010-7014 (2007).

155. Kizaki, N., Yasohara, Y. & Hasegawa, J. Carbonyl reductase, gene thereof and method of using the same. US 7,033,808 B2. (2006).

156. Kizaki, N.Y., Y.; Nagashima, N.; Hasegawa, J. Characterization of novel alcohol dehydrogenase of *Devosia riboflavina* involved in stereoselective reduction of 3-pyrrolidinone derivatives. *J Mol Cat B-Enzymatic* **51**, 73-80 (2008).

157. Mayer, K.M. & Arnold, F.H. A colorimetric assay to quantify dehydrogenase activity in crude cell lysates. *J Biomol Screen* **7**, 135-40 (2002).

158. Peters, M.W., Meinhold, P., Glieder, A. & Arnold, F.H. Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3. *J Am Chem Soc* **125**, 13442-13450 (2003).

159. Reetz, M.T. Directed evolution as a means to engineer enantioselective enzymes. in *Asymmetric Organic Synthesis with Enzymes* (eds. Gotor, V., Alfonso, I. & Garcia-Urdiales, E.) 21-56 (WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2008).

160. Schrader, W., Eipper, A., Pugh, D.J. & Reetz, M.T. Second-generation MS-based high-throughput screening system for enantioselective catalysts and biocatalysts. *Can J Chem* **80**, 626-632 (2002).

161. van Beilen, J.B., Kingma, J. & Witholt, B. Substrate specificity of the alkane hydroxylase system of *Pseudomonas oleovorans* GPo1. *Enz Microb Technol* **16**, 904-911 (1994).

162. Penning, T.M. & Jez, J.M. Enzyme redesign. *Chem Rev* **101**, 3027-46 (2001).

163. Harayama, S. Artificial evolution by DNA shuffling. *Trends Biotechnol* **16**, 76-82 (1998).

164. Zhao, H., Chockalingam, K. & Chen, Z. Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr Opin Biotechnol* **13**, 104-10. (2002).

165. Arnold, F.H. & Volkov, A.A. Directed evolution of biocatalysts. *Curr Opin Chem Biol* **3**, 54-9. (1999).

166. Feng, X. et al. The heme monooxygenase cytochrome P450cam can be engineered to oxidize ethane to ethanol. *Angew Chem Int Ed* **44**, 4029-4032 (2005).

167. Bartsch, S., Kourist, R. & Bornscheuer, U.T. Complete inversion of enantioselectivity towards acetylated tertiary alcohols by a double mutant of a *Bacillus subtilis* esterase. *Angew Chem Int Ed* **47**, 1508-11 (2008).

168. Ivancic, M., Valinger, G., Gruber, K. & Schwab, H. Inverting enantioselectivity of *Burkholderia gladioli* esterase EstB by directed and designed evolution. *J Biotechnol* **129**, 109-22 (2007).

169. Koga, Y., Kato, K., Nakano, H. & Yamane, T. Inverting enantioselectivity of *Burkholderia cepacia* KWI-56 lipase by combinatorial mutation and high-throughput screening using single-molecule PCR and in vitro expression. *J Mol Biol* **331**, 585-92 (2003).

170. Magnusson, A.O., Takwa, M., Hamberg, A. & Hult, K. An S-selective lipase was created by rational redesign and the enantioselectivity increased with temperature. *Angew Chem Int Ed* **44**, 4582-4585 (2005).

171. May, O., Nguyen, P.T. & Arnold, F.H. Inverting enantioselectivity by directed evolution of hydantoinase for improved production of l-methionine. *Nat Biotechnol* **18**, 317-320 (2000).

172. van Den Heuvel, R.H., Fraaije, M.W., Ferrer, M., Mattevi, A. & van Berkel, W.J. Inversion of stereospecificity of vanillyl-alcohol oxidase. *Proc Nat Acad Sci USA* **97**, 9455-60 (2000).

173. Reetz, M.T., Kahakeaw, D. & Sanchis, J. Shedding light on the efficacy of laboratory evolution based on iterative saturation mutagenesis. *Mol BioSyst* **5**, 115-122 (2009).

174. Bloom, J.D., Labthavikul, S.T., Otey, C.R. & Arnold, F.H. Protein stability promotes evolvability. *Proc Nat Acad Sci USA* **103**, 5869-74 (2006).

175. Yano, J.K. et al. Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus*. *J Biol Chem* **275**, 31086-31092 (2000).

176. Oku, Y. et al. Structure and direct electrochemistry of cytochrome P450 from the thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7. *J Inorg Biochem* **98**, 1194-1199 (2004).

177. Filho, M.V., Stillger, T., Müller, M., Liese, A. & Wandrey, C. Is log *P* a convenient criterion to guide the choice of solvents for biphasic enzymatic reactions? *Angew Chem Int Ed* **42**, 2993-2996 (2003).

178. Stampfer, W., Edegger, K., Kosjek, B., Faber, K. & Kroutil, W. Simple biocatalytic access to enantiopure (*S*)-1-heteroarylethanols employing a microbial hydrogen transfer reaction. *Adv Synth Catal* **346**, 57-62 (2004).

179. Stampfer, W., Kosjek, B., Moitzi, C., Kroutil, W. & Faber, K. Biocatalytic asymmetric hydrogen transfer. *Angew Chem Int Ed* **41**, 1014-7 (2002).

180. Gröger, H. et al. Preparative asymmetric reduction of ketones in a biphasic medium with an (S)-alcohol dehydrogenase under in situ-cofactor-recycling with a formate dehydrogenase. *Tetrahedron* **60**, 633-640 (2004).

181. Gröger, H. et al. Enantioselective reduction of 4-fluoroacetophenone at high substrate concentration using a tailor-made recombinant whole-cell catalyst. *Adv Synth Catal* **349**, 709-712 (2007).

182. Hummel, W., Abokitse, K., Drauz, K., Rollmann, C. & Gröger, H. Towards a large-scale asymmetric reduction process with isolated enzymes: Expression of an (*S*)-alcohol dehydrogenase in *E. coli* and studies on the synthetic potential of this biocatalyst. *Adv Synth Catal* **345**, 153-159 (2003).

183. Kataoka, M. et al. Stereoselective reduction of ethyl 4-chloro-3-oxobutanoate by Escherichia coli transformant cells coexpressing the aldehyde reductase and glucose dehydrogenase genes. *Appl Microbiol Biotechnol* **51**, 486-90 (1999).

184. Kizaki, N. et al. Synthesis of optically pure ethyl (*S*)-4-chloro-3-hydroxybutanoate by *Escherichia coli* transformant cells coexpressing the carbonyl reductase and glucose dehydrogenase genes. *Appl Microbiol Biotechnol* **55**, 590-5 (2001).

185. Shorrock, V.J., Chartrain, M. & Woodley, J.M. An alternative bioreactor concept for application of an isolated oxidoreductase for asymmetric ketone reduction. *Tetrahedron* **60**, 781-788 (2004).

186. Wong, C., Drueckhammer, D.G. & Sweers, H.M. Enzymatic vs. fermentative synthesis: Thermostable glucose dehydrogenase catalyzed regeneration of NAD(P)H for use in enzymatic synthesis. *J Am Chem Soc* **107**, 4028-4031 (1985).

187. Wong, C.-H. & Whitesides, G.M. Enzyme-catalyzed organic synthesis: NAD(P)H cofactor regeneration by using glucose-6-phosphate and the glucose-5-phosphate dehydrogenase from *Leuconostoc mesenteroides*. *J Am Chem Soc* **103**, 4890-4899 (1981).

188. Berg, J., Tymoczko, J. & Stryer, L. (eds.). *Biochemistry*, (W.H. Freeman and Company, New York, 2003).

189. Kupfer, D. & Holm, K.A. Prostaglandin metabolism by hepatic cytochrome P450. *Drug Metab Rev* **20**, 753-764 (1989).

190. el-Monem, A., el-Refai, H., Sallam, A.R. & Geith, H. Microbial 11 -hydroxylation of progesterone. *Acta Microbiol Pol B* **4**, 31-6 (1972).

191. Durst, F. & O'Keefe, D.P. Plant cytochromes P450: An overview. *Drug Metabol Drug Interact* **12**, 171-87 (1995).

192. Holton, T.A. et al. Cloning and expression of cytochrome P450 genes controlling flower colour. *Nature* **366**, 276-9 (1993).

193. Goldstein, J.A. & Faletto, M.B. Advances in mechanisms of activation and deactivation of environmental chemicals. *Environ Health Perspect* **100**, 169-76 (1993).

194. Munro, A.W. et al. P450 BM3: The very model of a modern flavocytochrome. *Trends Biochem Sciences* **27**, 250-257 (2002).

195. Budde, M., Maurer, S.C., Schmid, R.D. & Urlacher, V.B. Cloning, expression and characterisation of CYP102A2, a self-sufficient P450 monooxygenase from *Bacillus subtilis*. *Appl Environ Microbiol* **66**, 180-6 (2004).

196. Kitazume, T., Takaya, N., Nakayama, N. & Shoun, H. *Fusarium oxysporum* fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic counterpart of *Bacillus megaterium* cytochrome P450BM3. *J Biol Chem* **275**, 39734-39740 (2000).

197. Champion, P.M., Gunsalus, I.C. & Wagner, G.C. Resonance raman investigations of cytochrome P450CAM from *Pseudomonas putida*. *J Am Chem Soc* **100**, 3743-3751 (1978).

198. Takaya, N. et al. Cytochrome P450nor, a novel class of mitochondrial cytochrome P450 involved in nitrate respiration in the fungus *Fusarium oxysporum*. *Arch Biochem Biophys* **372**, 340-6 (1999).

199. Fruetel, J.A., Mackman, R.L., Peterson, J.A. & Ortiz de Montellano, P.R. Relationship of active site topology to substrate specificity for cytochrome P450terp (CYP108). *J Biol Chem* **269**, 28815-28821 (1994).

200. Kellner, D.G., Maves, S.A. & Sligar, S.G. Engineering cytochrome P450s for bioremediation. *Curr Opin Biotechnol* **8**, 274-8 (1997).

201. Chen, X. et al. Crystal structure of the F87W/Y96F/V247L mutant of cytochrome P450cam with 1,3,5-trichlorobenzene bound and further protein engineering for the oxidation of pentachlorobenzene and hexachlorobenzene. *J Biol Chem* **277**, 37519-26 (2002).

202. Narhi, L.O. & Fulco, A.J. Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P-450 monooxygenase induced by barbiturates in *Bacillus megaterium*. *J Biol Chem* **261**, 7160-9 (1986).

203. Wen, L.P. & Fulco, A.J. Cloning of the gene encoding a catalytically self-sufficient cytochrome P-450 fatty acid monooxygenase induced by barbiturates in *Bacillus megaterium* and its functional expression and regulation in heterologous (*Escherichia coli*) and homologous (*Bacillus megaterium*) hosts. *J Biol Chem* **262**, 6676-82 (1987).

204. Shono, T., Matsumura, Y., Uchida, K., Tsubata, K. & Makino, A. Electroorganic chemistry. 79. Efficient synthesis of pyrrolizidine and indolizidine alkaloids utilizing anodically prepared .alpha.-methyoxy carbamates as key intermediates. *J Org Chem* **49**, 300-304 (1984).

205. Reetz, M.T. & Carballeira, J.D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat Protocols* **2**, 891-903 (2007).

206. Wang, W., Xu, Y., Wang, D.I.C. & Li, Z. Recyclable Nanobiocatalyst for Enantioselective Sulfoxidation: Facile Fabrication and High Performance of Chloroperoxidase-Coated Magnetic Nanoparticles with Iron Oxide Core and Polymer Shell. *J Am Chem Soc* **131**, 12892-12893 (2009).

Appendix: Publications and Oral Presentations

Publications:

1. W. Tang, Z. Li, and H. Zhao. Inverting the Enantioselectivity of P450pyr Monooxygenase-Catalyzed Asymmetric Biohydroxylation by Directed Evolution. Chem Comm, 46, 5461-5463, 2010.

2. W. Tang, N. U. Nair, D. Eriksen, and H. Zhao. Industrial Applications of Enzymes as Catalysts. In Manual of Industrial Microbiology and Biotechnology, 3rd Ed. (A. L. Demain, R. Baltz, and J. E. Davies, Eds.), ASM Press, Washington, DC, 2010.

3. W. Tang and H. Zhao. Industrial Biotechnology: Tools and Applications. Biotechnol. J., 4, 1725-1739, 2009.

4. Y. Chen, W. Tang, J. Mou, and Z. Li. High-Throughput Method for Determining the Enantioselectivity of Enzyme-catalyzed Hydroxylations Based on Mass Spectrometry. Angew Chem Int Ed, 49, 5278-5283, 2010.

5. W. Zhang, W. Tang, Z. Wang, and Z. Li. Regio- and Stereo-selective Biohydroxylations with a Recombinant Escherichia coli expressing P450pyr Monooxygenase of Sphingomonas sp. HXN-200. Adv Syn Cat, 352, 3380-3390, 2010.

6. W. Zhang, W. Tang, D. Wang, and Z. Li, Concurrent Oxidations with Tandem Biocatalysts in One Pot: Green, Selective and Clean Oxidations of Methylene Groups to Ketones, Chem Comm, 47, 3284-3286, 2011.

Oral Presentations:

1. W. Tang, Z. Li and H. Zhao. Directed Evolution of an Enantioselective P450pyr for the Preparation of Chiral Pharmaceutical Intermediates. ACS National Meeting, Washington D.C. Aug 2009.

2. W. Tang, Z. Li and H. Zhao. Inverting the enantioselectivity of P450pyr Monooxygenase by Directed Evolution. ChemBioTech Conference, Singapore, Jan 2010.